



Health Technology Assessment

Volume 30 • Issue 56 • July 2026

ISSN 2046-4924

Surgery for early structural osteoarthritis of the hip and knee: the SCORE evidence synthesis and economic evaluation

*Andrew Price, Emma Searle, Maggie Westby, Gianluca Fabiano, Pamela Royle, Jill Colquitt,
Emma Loveman, Karen Barker, Sion Glyn-Jones, Antony Palmer, Hema Mistry, Toby Smith,
Andy Metcalfe, Rafael Pinedo-Villanueva and Sally Hopewell*

DOI: 10.3310/TNFD7352





Extended Research Article

Surgery for early structural osteoarthritis of the hip and knee: the SCORE evidence synthesis and economic evaluation

Andrew Price^{1*}, Emma Searle¹, Maggie Westby¹, Gianluca Fabiano¹,
Pamela Royle², Jill Colquitt³, Emma Loveman³, Karen Barker¹,
Sion Glyn-Jones¹, Antony Palmer¹, Hema Mistry², Toby Smith²,
Andy Metcalfe², Rafael Pinedo-Villanueva¹ and Sally Hopewell¹

¹Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, The Botnar Institute, Nuffield Orthopaedic Centre, Oxford, UK

²Warwick Clinical Trials Unit, Warwick Medical School, University of Warwick, Coventry, UK

³Southampton Health Technology Assessments Centre, Wessex Institute for Health Research and Development, University of Southampton, Southampton, UK

*Corresponding author andrew.price@ndorms.ox.ac.uk

Published July 2026
DOI: 10.3310/TNFD7352

This article should be referenced as follows:

Price A, Searle E, Westby M, Fabiano G, Royle P, Colquitt J, *et al.* Surgery for early structural osteoarthritis of the hip and knee: the SCORE evidence synthesis and economic evaluation. *Health Technol Assess* 2026;**30**(56). <https://doi.org/10.3310/TNFD7352>

Health Technology Assessment

ISSN 2046-4924 (Online)

Impact factor: 4

A list of Journals Library editors can be found on the [NIHR Journals Library website](#)

Launched in 1997, *Health Technology Assessment* (HTA) has an impact factor of 4 and is ranked 30th (out of 174 titles) in the 'Health Care Sciences & Services' category of the Clarivate 2022 Journal Citation Reports (Science Edition). It is also indexed by MEDLINE, CINAHL (EBSCO Information Services, Ipswich, MA, USA), EMBASE (Elsevier, Amsterdam, the Netherlands), NCBI Bookshelf, DOAJ, Europe PMC, the Cochrane Library (John Wiley & Sons, Inc., Hoboken, NJ, USA), INAHTA, the British Nursing Index (ProQuest LLC, Ann Arbor, MI, USA), Ulrichsweb™ (ProQuest LLC, Ann Arbor, MI, USA) and the Science Citation Index Expanded™ (Clarivate™, Philadelphia, PA, USA).

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta.

Criteria for inclusion in the *Health Technology Assessment* journal

Manuscripts are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

This article

The research reported in this issue of the journal was commissioned and funded by the HTA programme as award number NIHR127398. The protocol was agreed in June 2019. The draft manuscript began editorial review in May 2025 and was accepted for publication in November 2025. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' manuscript and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this article.

This article presents independent research funded by the National Institute for Health and Care Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, the HTA programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, the HTA programme or the Department of Health and Social Care.

This article was published based on current knowledge at the time and date of publication. NIHR is committed to being inclusive and will continually monitor best practice and guidance in relation to terminology and language to ensure that we remain relevant to our stakeholders.

Copyright © 2026 Price *et al.* This work was produced by Price *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: <https://creativecommons.org/licenses/by/4.0/>. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Newgen Digitalworks Pvt Ltd, Chennai, India (www.newgen.co).

Abstract

Research question: Is surgery for early hip or knee osteoarthritis clinically effective and cost-effective?

Background: Osteoarthritis is the most common joint disease worldwide costing between 1% and 2.5% of gross domestic product in developed countries. Once a hip or knee develops end-stage disease, arthroplasty surgery is highly successful in most patients. Treatment options are limited for early disease, and patients may fall into a 'treatment gap', despite having severe symptoms. Many surgical procedures are used to treat early osteoarthritis of the hip and knee, but the clinical and cost-effectiveness of these techniques remains unclear.

Overall aim: To investigate the role of surgery in the management of early hip or knee osteoarthritis?

Objectives:

1. Use systematic review and evidence synthesis to determine the clinical effectiveness of current surgical interventions for early hip and knee osteoarthritis.
2. Estimate the cost-effectiveness of these procedures using decision model analytic modelling.
3. Create clinical treatment guidance and identify areas where new clinical research is required.

Methods: The programme of work consisted of systematic review, evidence synthesis and health economic modelling for surgical treatments used for early hip and knee osteoarthritis. We divided treatments into four groups: (1) non-regenerative procedures, (2) regenerative procedures, (3) load-modifying procedures and (4) arthroplasty procedures. They were compared with non-surgical treatment following National Institute for Health and Care Excellence guidance. We used outcome measures that satisfied the Outcome Measures in Rheumatoid Arthritis Clinical Trials and Osteoarthritis Research Society International core outcome set for hip and knee osteoarthritis. Decision-analytic models were developed to assess the cost-effectiveness of treatment alternatives expressed as cost per quality-adjusted life-year gained.

Results: The systematic review identified 61,098 studies, and our evidence synthesis identified a number of surgical techniques that have been used to treat early hip and knee osteoarthritis. However, we found very little meaningful evidence of the clinical efficacy for the majority of the surgical procedures. We only found seven randomised controlled trials and three non-randomised studies that compared a surgical technique to non-operative treatment. There were two surgical treatments that had limited but supportive evidence. Femoro-acetabular impingement is an early stage of hip osteoarthritis, and there is some evidence that arthroscopic chondroplasty provides patients with clinical relief of symptoms, above that seen with non-operative measures. In the knee, weaker evidence suggests arthroscopic debridement of osteoarthritic articular cartilage fragments may be of benefit, above non-operative measures. Health economic analysis suggests that both procedures may be cost-effective, but the evidence is again of low quality.

Limitations: The original updated searches completed in January 2021, with a further search for randomised controlled studies completed in January 2025.

Conclusions: There is a significant gap in the evidence base supporting surgical interventions in early hip and knee osteoarthritis. Limited evidence suggests that arthroscopic chondroplasty procedures may be effective in treating hip femoro-acetabular impingement and early knee osteoarthritis. A small number of randomised studies are in progress, but these studies do not encompass all procedures that are in use.

Future work: The key recommendation for future work is that more high-quality comparative studies are required to establish if the surgical interventions that are currently used are effective.

Study registration: This study is registered as PROSPERO CRD42019156444.

Funding: This award was funded by the National Institute for Health and Care Research (NIHR) Health Technology Assessment programme (NIHR award ref: NIHR127398) and is published in full in *Health Technology Assessment*; Vol. 30, No. 56. See the NIHR Funding and Awards website for further award information.

Contents

List of tables	ix
List of figures	xi
List of abbreviations	xii
Plain language summary	xiv
Scientific summary	xv
Chapter 1 Background and rationale	1
Describing the condition	1
<i>Current practice (National Institute for Health and Care Excellence osteoarthritis guideline)</i>	1
Health Technology Assessment project	1
Aims of the project	2
Key objectives	2
Definition of early osteoarthritis of the knee and hip	2
Surgical treatment for early hip and knee osteoarthritis	4
<i>Joint preservation: non-regenerative procedures</i>	4
<i>Joint preservation: biological reconstruction of the joint</i>	4
<i>Joint preservation: load-modifying procedures</i>	5
<i>Joint replacement procedures</i>	5
Chapter 2 Evidence synthesis methods	7
Evidence synthesis protocol	7
Criteria for inclusion and exclusion of studies	7
<i>Types of studies</i>	7
<i>Setting</i>	8
<i>Target population</i>	8
<i>Exclusion criteria</i>	10
Interventions	10
<i>Joint preservation: non-regenerative procedures</i>	10
<i>Joint preservation: biological reconstruction</i>	11
<i>Joint preservation: load-modifying procedures</i>	12
<i>Joint replacement procedures</i>	12
Comparator interventions	13
Outcomes	14
Minimum duration of follow-up	16
Search strategy	16
Data collection and analysis	17
<i>Selection of studies</i>	17
<i>Documenting exclusions</i>	17
<i>Data extraction and management</i>	17
Assessment of risk of bias in included studies	18
Main planned comparisons	19
Data synthesis and measures of treatment effect	19
Dealing with missing data	20
Assessment of reporting biases	20

Subgroup analysis and investigation of heterogeneity	20
Sensitivity analysis	21
Summary of findings tables	21
Overviews of reviews of non-comparative studies to inform research	21
Data to be taken forward to the health economic analysis	21
Time-line and adjustments to protocol	21
Chapter 3 Non-regenerative procedures used to treat early osteoarthritis in the knee	22
Results of the search	22
Included studies characteristics	23
<i>Study details</i>	23
<i>Characteristics of included studies</i>	23
<i>Population, intervention, control/comparison, outcome and participant characteristics</i>	23
<i>Surgical interventions assessed</i>	23
<i>Comparator interventions assessed</i>	28
<i>Outcomes</i>	28
Results	29
<i>Direct treatment of early osteoarthritis</i>	29
<i>Indirect treatment of early osteoarthritis</i>	35
Conclusions	41
Chapter 4 Non-regenerative surgical procedures used to treat early osteoarthritis in the hip	42
Results of the search	42
<i>Ongoing studies characteristics</i>	42
Included studies characteristics	42
<i>Study details</i>	42
<i>Characteristics of included studies</i>	42
<i>Population, intervention, control/comparison, outcome characteristics</i>	44
Results	44
<i>Type 1: comparison of non-regenerative surgical procedures versus non-operative management</i>	47
<i>Type 2: comparison of non-regenerative surgical procedure versus another non-regenerative surgery</i>	51
<i>Type 3: comparison of non-regenerative surgical procedure versus another surgical procedure</i>	51
Conclusions	51
Chapter 5 Load-modifying procedures to treat early osteoarthritis in the knee	52
Results of the search	52
Included studies characteristics	52
<i>Characteristics of included studies</i>	52
<i>Population, intervention, control/comparison, outcome characteristics</i>	55
<i>Outcomes</i>	55
Results	55
<i>Type 1 studies: comparison of load-modifying surgical procedures versus non-operative management</i>	55
<i>Type 2 studies: comparison of one load-modifying surgical procedure versus another load-modifying procedure or other joint-preserving surgical procedures</i>	56
<i>Type 3 studies: comparison of load-modifying surgical procedure versus joint replacement surgery</i>	57
Conclusions	57
Chapter 6 Load-modifying surgical procedures around used to treat early hip osteoarthritis	58
Results of the search	58
Included studies characteristics	58
<i>Study details</i>	58
<i>Characteristics of included studies</i>	59

Population, intervention, control/comparison, outcome characteristics	60
Comparator interventions assessed	60
Outcomes	61
Results	61
Type 1: comparison of load-modifying surgical procedures versus non-operative management	61
Type 2: comparison of load-modifying surgical procedures versus another load-modifying procedure	61
Type 3: comparison of load-modifying surgical procedures versus another surgical procedure	61
Additional analysis	61
Conclusions	66
Chapter 7 Biological surgical reconstruction for early knee osteoarthritis	67
Results of the search	67
Included studies characteristics	67
Study details	67
Characteristics of included studies	67
Population, intervention, control/comparison, outcome	67
Surgical interventions assessed	67
Comparator interventions assessed	67
Outcomes	67
Results	67
Type 1: comparisons of biological surgical reconstruction versus non-operative management for treatment of patients with early knee osteoarthritis	68
Type 2: comparisons of biological surgical reconstruction versus another biological surgical reconstruction procedure for treatment of patients with early knee osteoarthritis	69
Type 3: comparisons of biological surgical reconstruction procedures versus another surgical options for treatment of patients with early knee osteoarthritis	69
Additional analysis	69
Conclusions	72
Chapter 8 Biological surgical reconstruction for early hip osteoarthritis	73
Results of the search	73
Included studies characteristics	73
Additional studies of interest included	73
Studies included	73
Outcomes	74
Results	75
Microfracture studies	75
Comparison of autologous matrix-induced chondrogenesis and microfracture	75
Comparison of autologous matrix-induced chondrogenesis and matrix-associated autologous chondrocyte implantation	78
Bone marrow aspirate use in patients undergoing non-regenerative hip surgery for femoro-acetabular impingement	78
Conclusions	78
Chapter 9 Joint replacement used to treat early knee osteoarthritis	79
Results of the search	79
Included studies characteristics	79
Study details	79
Characteristics of included studies	79
Population, intervention, control/comparison, outcome characteristics	79
Surgical interventions assessed	79
Comparator interventions assessed	82
Outcomes	82

Results	82
<i>Direct treatment of early knee osteoarthritis</i>	83
<i>Indirect treatment of early knee osteoarthritis</i>	83
Additional analysis	84
Conclusions	88
Chapter 10 Joint replacement performed for early hip osteoarthritis	90
Results of the search	90
Included studies characteristics	90
Additional analysis of studies outside the search criteria	90
<i>Studies included</i>	90
<i>Prognostic factors assessed</i>	92
<i>Outcomes assessed</i>	92
Results	94
Conclusions	95
Chapter 11 Cost-effectiveness of surgical interventions for early knee and hip osteoarthritis	96
Introduction	96
Model structure	96
Model inputs	97
<i>Transition probabilities</i>	97
<i>Health-related quality of life</i>	100
<i>Resource use and costs</i>	101
<i>Complications</i>	106
<i>Mortality</i>	106
Sensitivity and scenario analyses	108
<i>Knee</i>	108
<i>Hip</i>	108
Results	109
<i>Knee</i>	110
<i>Hip</i>	113
Discussion and conclusions	113
Chapter 12 Discussion and research recommendations	119
Introduction	119
Definition of early osteoarthritis	119
Comparator interventions	119
Levels of evidence	119
Types of surgical procedure	120
Surgical procedures for treatment of early knee osteoarthritis	120
<i>Non-regenerative surgical procedures for treatment of early knee osteoarthritis</i>	120
<i>Load-modifying surgical procedures for treatment of early knee osteoarthritis</i>	120
<i>Biological reconstruction surgical procedures for treatment of early knee osteoarthritis</i>	120
<i>Joint replacement surgical procedures for treatment of early knee osteoarthritis</i>	121
Surgical procedures for treatment of early hip osteoarthritis	121
<i>Non-regenerative surgical procedures for treatment of early hip osteoarthritis</i>	121
<i>Load-modifying surgical procedures for treatment of early hip osteoarthritis</i>	121
<i>Biological reconstruction surgical procedures for treatment of early hip osteoarthritis</i>	121
<i>Joint replacement surgical procedures for treatment of early hip osteoarthritis</i>	121
Patient and public involvement	122
Equality, diversity and inclusion	122
Limitations of research	122
Research recommendations	123

CONTENTS

Additional information	124
References	126
Appendix 1	138
Appendix 2	144

List of tables

TABLE 1 Outline diagnostic criteria for early structural OA of the hip and knee	3
TABLE 2 Summary of included non-regenerative surgery studies characteristics for direct treatment of early knee OA	24
TABLE 3 Summary of included non-regenerative surgery studies characteristics for indirect treatment of early knee OA	26
TABLE 4 Summary of findings table for non-regenerative surgery studies for treatment of early knee OA	30
TABLE 5 Summary of included non-regenerative surgery studies characteristics for treatment of early hip OA	45
TABLE 6 Summary of findings table for non-regenerative surgery studies for treatment of early hip OA	48
TABLE 7 Summary of included load-modifying surgery studies characteristics for treatment of early knee OA	54
TABLE 8 Summary of findings table for load-modifying surgery studies for treatment of early knee OA	56
TABLE 9 Summary of included load-modifying surgery studies characteristics for treatment of early hip OA	60
TABLE 10 Summary of additionally included studies of load-modifying surgery for treatment of early hip OA	62
TABLE 11 Summary of included biological reconstruction surgery studies characteristics for treatment of early knee OA	69
TABLE 12 Summary of additional comparative studies of biological interventions in the knee	70
TABLE 13 Summary of included biological reconstruction surgery studies characteristics for treatment of early hip OA	75
TABLE 14 Summary of additional comparative studies of biological interventions in the hip	76
TABLE 15 Summary of included joint arthroplasty surgery studies characteristics for direct treatment of early knee OA	81
TABLE 16 Summary of included joint arthroplasty surgery studies characteristics for indirect treatment of early knee OA	82
TABLE 17 Summary of additional prognostic studies of interest where joint arthroplasty surgery has been used in the treatment of early knee OA	85
TABLE 18 Summary of included joint arthroplasty surgery studies characteristics for treatment of early hip OA	92
TABLE 19 Summary of additional prognostic studies of interest where joint arthroplasty surgery has been used in the treatment of early hip OA	93
TABLE 20 Transitions probabilities – knee model	100
TABLE 21 Transitions probabilities – hip model	100

TABLE 22 Utility inputs – knee model	103
TABLE 23 Cost-effectiveness results for DSA model	103
TABLE 24 Utility estimates – hip model	105
TABLE 25 Cost inputs – knee model	105
TABLE 26 Cost inputs – hip model	107
TABLE 27 Probabilities of transition to death for both knee and hip models	107
TABLE 28 Probability distributions for transitions – knee model	109
TABLE 29 Probability distributions for QALYs – knee model	109
TABLE 30 Probability distributions for costs – knee model	110
TABLE 31 Probability distributions for transitions – hip model	112
TABLE 32 Probability distributions for QALYs – hip model	112
TABLE 33 Probability distributions for costs – hip model	112
TABLE 34 Quality-adjusted life-year inputs for DSA – hip model	113
TABLE 35 Cost-effectiveness result for the hip model	115
TABLE 36 Cost-effectiveness result for the knee model	117
TABLE 37 Risk of bias in non-randomised comparative studies for non-regenerative procedures used to treat early knee OA	146
TABLE 38 Risk of bias in non-randomised comparative studies of non-regenerative procedures used to treat early hip OA	147
TABLE 39 Risk of bias in non-randomised comparative studies involving load-modifying procedures used to treat early knee OA	149

List of figures

FIGURE 1	PRISMA diagram: non-regenerative surgery for treatment of early knee OA	22
FIGURE 2	PRISMA diagram: non-regenerative surgery for treatment of early hip OA	43
FIGURE 3	PRISMA diagram: load-modifying surgery for treatment of early knee OA	53
FIGURE 4	PRISMA diagram: load-modifying surgery for treatment of early hip OA	59
FIGURE 5	PRISMA diagram: biological reconstruction surgery for treatment of early knee OA	68
FIGURE 6	PRISMA diagram: biological reconstruction surgery for treatment of early hip OA	74
FIGURE 7	PRISMA diagram: knee arthroplasty surgery for treatment of early knee	80
FIGURE 8	PRISMA diagram: knee arthroplasty surgery for treatment of early hip OA	91
FIGURE 9	Markov model for knee and hip analyses	98
FIGURE 10	Annual transitions to knee joint replacement	99
FIGURE 11	(a) Mapped EQ-5D from RCT study and (b) EQ-5Ds after matching baseline values – knee model	102
FIGURE 12	A diagram showing the utility inputs for the knee model	102
FIGURE 13	Mapped EQ-5D and utility inputs – hip model	104
FIGURE 14	A diagram showing the utility inputs for the hip model	105
FIGURE 15	A figure showing the cost inputs for the knee model	106
FIGURE 16	A figure showing the cost inputs for the hip model	107
FIGURE 17	Utility estimates for intervention/post-intervention states under each scenario – knee model	111
FIGURE 18	Figure showing QALY inputs for DSA for the hip model	113
FIGURE 19	Cost-effectiveness plane (a) and CEAC (b) – knee model	114
FIGURE 20	Cost-effectiveness acceptability curve base case vs. scenarios	115
FIGURE 21	Cost-effectiveness plane (a) and CEAC (b) – hip model	116
FIGURE 22	Cost-effectiveness plane (a) and CEAC (b) – DSA hip model	118
FIGURE 23	Risk of bias in RCTs for non-regenerative procedures used to treat early knee OA	144
FIGURE 24	Risk of bias in RCTs of non-regenerative procedures used to treat early hip OA	146
FIGURE 25	Risk of bias in RCTs involving load-modifying procedures used to treat early knee OA	148

List of abbreviations

ACI	autologous chondrocyte implantation	HTO	high tibial osteotomy
ADL	activities of daily living	IA	intra-articular
AE	adverse event	ICER	incremental cost-effectiveness ratio
AIC	Akaike information criterion	ICRS	International Cartilage Repair Society
AKSS	American Knee Society Score	iHOT	International Hip Outcome Tool
AMIC	autologous matrix-induced chondrogenesis	IKDC	International Knee Documentation Committee
APM	arthroscopic partial meniscectomy	INAHTA	International Network of Agencies for Health technology Assessment
AUC	area under the curve	IPD	individual participant data
BIC	Bayesian information criterion	IQWiG	Institute for Quality and Efficiency in Health Care
BMA	bone marrow aspirate	JSN	joint space narrowing
BMC	bone marrow concentrate	JSW	joint space width
BMI	body mass index	KJD	knee joint distraction
CEAC	cost-effectiveness acceptability curve	KL	Kellgren-Lawrence
DFO	distal femoral osteotomy	KOOS	Knee Injury and Osteoarthritis Outcome Score
DSA	deterministic sensitivity analysis	KSPS	Knee Specific Pain Scale
EKOA	early knee osteoarthritis	KSS	Knee Society Score
EHOA	early hip osteoarthritis	MACI	matrix-associated autologous chondrocyte implantation
EMBASE	Excerpta Medica dataBASE	MCID	minimum clinically important difference
EQ-5D	EuroQol-5 Dimensions	MF	microfracture
FAI	femoro-acetabular impingement	mHHS	modified Harris Hip Score
GRADE	Grading of Recommendations Assessment, Development and Evaluation	MRI	magnetic resonance imaging
FTCL	full-thickness cartilage loss	MSD	mechanical shaver debridement
HA	hyaluronic acid	NICE	National Institute for Health and Care Excellence
HAGOS	Copenhagen Hip and Groin Outcome Score	NIHR	National Institute for Health and Care Research
HCHS	Hospital and Community Health Service	NRS	non-randomised study
HHS	Harris Hip Score	NSAID	non-steroidal anti-inflammatory drug
HOOS	Hip Disability and Osteoarthritis Outcome Score	OA	osteoarthritis
HOS	Hip Outcome Score	OARSI	Osteoarthritis Research Society International
HRQoL	health-related quality of life	OCA	osteochondral allograft
HSS	Hospital for Special Surgery		
HTA	Health Technology Assessment		

OHS	Oxford Hip Score	SA	sensitivity analysis
OKS	Oxford Knee Score	SAE	serious adverse event
OMERACT	Outcome Measures in Rheumatoid Arthritis Clinical Trials	SF-12	Short Form questionnaire-12 items
PAO	peri-acetabular osteotomy	SF-36	Short Form questionnaire-36 items
PICO	population, intervention, control/ comparison, outcome	SIGN	Scottish Intercollegiate Guidelines Network
PPI	patient and public involvement	SMD	standardised mean difference
PREM	patient-reported experience measure	THR	total hip replacement
PRISMA	Preferred Reporting Items for Systematic review and Meta-Analyses	THA	total hip arthroplasty
PROM	patient-reported outcome measure	TKA	total knee arthroplasty
PSA	probabilistic sensitivity analysis	TKR	total knee replacement
PTCL	partial thickness cartilage loss	UC	usual care
QALY	quality-adjusted life-year	UKA	unicompartmental knee arthroplasty
QoL	quality of life	UKR	unicompartmental knee replacement
q-RCT	quasi-randomised controlled trial	VAS	visual analogue scale
RAO	rotational acetabular osteotomy	WOMAC®	Western Ontario and McMaster Universities Arthritis Index
RCT	randomised controlled trial		
RFC	radiofrequency chondroplasty	WTP	willingness to pay

Plain language summary

Osteoarthritis of the hip or knee joint is a common problem. As the joint is progressively damaged, people affected develop pain and stiffness. For many, simple measures, such as exercise, weight loss and medication, can control the symptoms. For others, symptoms remain severe, normal daily activities are difficult and some cannot work. When joint damage is less severe, described as 'early osteoarthritis', joint replacement is thought to be less effective and different surgical treatments are used. However, currently, it is not clear which procedures are most useful in helping relieve patient's symptoms. Patients could be potentially missing out on surgery that is beneficial, while others may undergo treatments that have become established but are not effective. Therefore, the aim of this research was to summarise the existing evidence for surgical treatments to treat early hip and knee osteoarthritis. We have undertaken an extensive search for all published research studies that details the outcome after surgical treatments for early hip or knee osteoarthritis, focusing on studies that compare surgery to established non-operative treatment. We looked at pain relief, return to activity and the complications associated with each surgery. The large body of evidence collected was sorted and summarised to assess the success or failure of each type of surgery. Hip impingement represents a very early stage of osteoarthritis and there is some good evidence that keyhole shaving of abnormal tissue provides patients with clinical relief of symptoms, above that seen with non-operative measures. In the knee weaker evidence suggests key-hole surgical removal of articular cartilage fragments may be of benefit. Health economic analysis suggest that both procedures may be cost-effective, but the evidence is of low quality. We found very limited evidence to support the majority of the surgical techniques that are used to treat early hip and knee osteoarthritis. We recommend that further studies are conducted to fill this evidence gap.

Scientific summary

Background

Osteoarthritis (OA) is the most common joint disease worldwide costing between 1.0% and 2.5% of gross domestic product in developed countries. OA is a joint wide condition characterised by damage to the articular surfaces of the joint, bone sclerosis and cysts, stiffness of the soft tissues around the knee, synovitis and intermittent joint effusions. Many patients with structural hip and knee OA have no or very mild symptoms from their joint. However, there are a group of patients with early or late OA structural changes who develop more pronounced hip or knee pain and loss of function, described as clinical OA, which includes a broad set of symptoms, characterised by pain and stiffness in the joint that produces a reduction in function and ultimately in quality of life (QoL). All patients who present with clinical OA should undergo initial treatment following the National Institute for Health and Care Excellence (NICE) clinical guideline CG226, with good education on their condition and treatment options; therapeutic exercise; and weight loss if appropriate. Second-line additional treatments can be introduced, such as topical or oral non-steroidal anti-inflammatory drugs (NSAIDs) and third-line adjunctive measures can then be considered, including intra-articular (IA) injections, manual therapy and bracing. However, not all patients respond to non-operative treatment and some develop persistent severe symptoms. For patients in this group who have end-stage (late) structural hip or knee OA, joint replacement (arthroplasty) is a highly successful treatment option. However, there are patients with early structural disease with symptoms as severe as for end-stage disease, that are also unresponsive to non-operative treatment. Joint replacement is typically not offered to patients with earlier structural disease as it is believed to be less effective and these patients may fall into a 'treatment gap'. Many surgical procedures have been adopted to fill this gap in treatment for early hip or knee OA, but often with limited success and at significant expense. It remains unclear how and when these techniques are best employed, and a more evidence-based approach is required.

Aim of this study

To investigate the role for surgery in the management of early hip or knee OA.

Objectives

1. Use systematic review and evidence synthesis to determine the clinical effectiveness of current surgical interventions for early hip and knee OA.
2. Estimate the cost-effectiveness of these procedures using decision-analytic modelling.
3. Create clinical treatment guidance and identify areas where new clinical research is required

Methods

The programme of work consisted of systematic review, evidence synthesis and health economic modelling for surgical treatments used for early hip and knee OA.

Definition of early osteoarthritis

We have defined early OA as (1) the presence of typical OA symptoms and (2) radiological evidence of structural early OA and (3) the absence of widespread disease. Diagnostic criteria for early structural hip or knee OA were one or more of; (1) loss of joint space on a weight-bearing radiograph (excluding bone-on-bone change), (2) partial-thickness cartilage loss, cartilage delamination or focal full-thickness cartilage on magnetic resonance imaging, (3) cartilage changes affecting the joint seen at arthroscopy or open surgery. Painful hip dysplasia and femoro-acetabular impingement (FAI) of the hip were included, as they represent an initial phase of the hip OA.

Surgical interventions

We divided surgical treatments into four groups: (1) non-regenerative procedures (e.g. osteochondroplasty), (2) regenerative procedures [e.g. autologous chondrocyte implantation (ACI)], (3) load-modifying procedures (e.g. tibial or peri-acetabular osteotomy) and (4) arthroplasty procedures, including small implants.

Comparator treatment

The study focused on published evidence where surgical procedures were compared with non-surgical treatment. Non-surgical treatment was defined as measures outlined in NICE CG226 OA treatment guidance (e.g. patient education, therapeutic exercise, weight loss, NSAIDs, IA injections manual therapy and bracing).

Outcomes assessed

We used outcome measures that satisfied the Outcome Measures in Rheumatoid Arthritis Clinical Trials (OMERACT) and Osteoarthritis Research Society International core outcome set for hip and knee arthritis (and OMERACT Core Outcome Domain Set for arthroplasty), with clinical outcomes, including patient experience, radiological outcomes, and health economic outcomes with employment-related measures.

Evidence synthesis

The main search strategy included comparative studies of surgery and non-operative treatment including; randomised controlled trials (RCTs) or systematic reviews of these study designs, comparative non-randomised studies (NRS) and systematic reviews of these designs and systematic reviews of non-comparative study designs. Where there was a lack of comparative study evidence we included case series that offered prognostic insight into treatment. Eight broad comprehensive searches were completed covering the four surgical treatment groups for both hip and knee. The searches used MEDical Literature Analysis and Retrieval System (MEDLINE), Excerpta Medica dataBASE (EMBASE), The Cochrane Library, National Health Service Economic Evaluation Database, Health Technology Assessment (HTA) database, NICE database, International Network of Agencies for Health Technology Assessment (INAHTA) database, Scottish Intercollegiate Guidelines Network (SIGN) database, the Canadian Agency for Drugs and Technologies in Health, Institute for Quality and Efficiency in Health Care (IQWiG) database and Web of Science. In addition, the search included 'research in progress' through ClinicalTrials.gov, World Health Organization Clinical Trials Search Portal, ISRCTN.com and the National Institute for Health and Care Research awards database. Where comparative evidence existed, that met our inclusion criteria, we summarised the findings focusing on the mean difference or standardised mean difference in continuous clinical outcome data. Where multiple studies existed, we performed meta-analysis. We applied the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system to assess the certainty of the results to help in interpretation. Risk of bias was assessed for RCTs using the Cochrane Risk of Bias tool, applied to the main outcome measures of pain and function. For risk of bias in comparative NRS we will use tools developed by the National Institute for Health, National Heart, Lung and Blood Institute.

Health economic analysis

Two cohort state-transition 'Markov' models were developed to reflect the clinical pathways of patients with early hip or knee OA. These decision-analytic (cost-utility) models were used to estimate the cost-effectiveness of surgical versus non-operative management procedures expressed as cost per quality-adjusted life-year (QALY) gained, based on the clinical evidence available from the systematic review evidence where available.

Evidence synthesis results

In general, there was limited evidence that compared surgical procedures aimed at treating early knee or hip OA with non-operative measures.

Non-regenerative procedures

The majority of identified eligible evidence that met our criteria lay within this category of procedures. In the knee there were five comparative studies investigating the efficacy of arthroscopic articular cartilage debridement (four RCTs and one NRS) to directly treat early knee OA, where the summary of findings analysis demonstrated a positive clinical effect of surgery. In the hip there were four comparative studies investigating arthroscopic osteochondroplasty (three RCTs and one NRS) to directly treat FAI, where the summary of findings analysis demonstrated a positive clinical effect of surgery. However, in both knee and hip studies, GRADE analysis showed the evidence was of low certainty. In the

knee there were 19 RCTs of arthroscopic meniscectomy used to treat patients with meniscal tears, but they could not considered direct treatment of early knee OA. In the hip we identified a number of ongoing studies focused on labral repair as an alternative surgical treatment for FAI.

Load-modifying procedures

In the knee, there was one NRS of high-tibial osteotomy (HTO) compared with non-operative OA treatment. In addition, there was one RCT of two types of HTO. The summary of findings showed moderate benefit of surgery favouring surgery compared to non-operative treatment, but the evidence is of very low certainty. We identified one ongoing RCT of HTO compared to non-operative treatment. In the hip, our search identified cohorts studies, where peri-acetabular osteotomy has been used as an intervention used to treat early hip OA (dysplasia), with reports of clinical benefit to patients. There were no studies that had reported comparing its use to non-operative treatment, but we identified one RCT in progress.

Regenerative procedures

We identified a number of clinical cohort studies that show regenerative surgical techniques, such as microfracture (MF), mosaicplasty, osteochondral allografts, autologous matrix-induced chondrogenesis (AMIC) and ACI have been used to treat patients with early hip and knee OA. However, we identified no comparative studies of biological reconstruction surgery compared against non-operative treatment in an early OA group. There was one knee study that compared MF surgery against AMIC surgery, but no similar studies in the hip literature. More extensive evidence exists to support the use of ACI to treat isolated cartilage defects of the knee.

Arthroplasty procedures

We did not identify any studies where joint arthroplasty (with either total, partial or small implant procedures) had been compared with non-operative treatment in patients with early hip or knee OA. There has been one RCT of knee replacement compared with physiotherapy, in patients with generally more extensive OA changes. In this study, there were a small number of patients who met out criteria for early knee OA, but the limited no meaningful conclusions could be drawn from the work. We did identify a number of prognostic cohort studies that show hip arthroplasty is performed in patients who have early hip OA. These prognostic studies compare independent associations of early OA (vs. late OA) with patient-reported outcome measures, and we found there was weak evidence that the outcome of the intervention in the early OA group are worse than when compared to late-stage hip OA Kellgren-Lawrence grade. This evidence is mainly of very low certainty, but suggests that early OA may be associated with worse pain scores, poorer hip function, especially at short-term, medium-term and long-term follow-up.

Health economic analysis results

Evidence synthesis of the clinical effectiveness identified the following surgical procedures used in the treatment of early hip and knee OA, to show potentially sufficient evidence for their cost-effectiveness to be assessed: (1) arthroscopic debridement of early OA in the knee and (2) arthroscopic osteochondroplasty of FAI in the hip. In each case the models produced were developed over a time horizon of 5 years given the lack of evidence on the long-term costs and outcomes of patients receiving procedures for early OA. In the case of the hip analysis, our deterministic model found that arthroscopic debridement was expected to be cost-effective compared to non-operative management with an incremental cost-effectiveness ratio (ICER) of £14,602 per QALY gained. In the knee analysis, the discounted deterministic results showed that arthroscopic debridement would be expected to be cost-effective compared to non-operative management at an ICER of £5649 per QALY gained. In summary, for both the knee and hip analyses, it appears possible that arthroscopic debridement could be cost-effective compared to non-operative management. However, expected QoL gains do not reach minimally important differences and there is considerable uncertainty around the cost-effectiveness results; hence, more high-quality evidence is needed for a more conclusive cost-effectiveness finding to be produced.

Conclusions

Overall, our study identified a significant gap in the evidence base to support surgical interventions in early hip and knee OA, although many surgical techniques are used to treat this group of patients.

There were only two surgical techniques where we could demonstrate evidence to suggest that they are effective treatments, one in hip and one in knee patients. In the hip there is evidence of the efficacy of osteochondroplasty to treat hip FAI compared to non-operative treatment. There appears to be a modest additional benefit to surgery improving pain and function, leading to a potential cost-effectiveness benefit over a 5-year window. In similar fashion there is limited evidence suggesting that arthroscopic chondroplasty may be effective in treating early knee OA. But this is underpinned with clinical outcome data of low certainty, from very few studies. In both cases, results must be treated with some caution as both clinical and health economic results have a low-level of evidence certainty. A small number of randomised studies are in progress, including assessments of HTO in the knee and labral tear repair in the hip. The results of these studies will be useful additions to the evidence base, but do not encompass all procedures that are in use. A particular note must be made of the lack of studies relating to biological reconstruction surgery, suggesting that surgical techniques that are successful in treating isolated cartilage defects in the hip and knee joint, have not been translated into treatment of early OA. Similarly joint replacement, which is very established as a frontline treatment for late-stage hip and knee OA, has not been assessed as a treatment for early knee OA. This probably reflects prognostic cohort studies suggesting that a diagnosis of early hip and knee OA has a negative effect on clinical outcome for patients.

Our key recommendation is that more high-quality comparative randomised studies are required to establish if the surgical interventions that are currently used are effective when compared to non-operative treatments. In developing studies, methodology must include agreed widely used definitions of early OA and the non-operative treatments used to build consistency across studies.

Study registration

This study is registered as PROSPERO CRD42019156444.

Funding

This award was funded by the National Institute for Health and Care Research (NIHR) Health Technology Assessment programme (NIHR award ref: NIHR127398) and is published in full in *Health Technology Assessment*; Vol. 30, No. 56. See the NIHR Funding and Awards website for further award information.

Chapter 1 Background and rationale

Describing the condition

Worldwide musculoskeletal disorders accounted for 6–8% of all disability-adjusted life-years (DALYs) and osteoarthritis (OA) accounts for approximately 10% of this disease burden.¹ In the UK, more people receive disability living allowance as a result of arthritis than received for heart disease, stroke, chest disease and cancer combined (Department of Work and Pensions). OA affecting the knee is common with a global prevalence of radiographically confirmed symptomatic knee OA estimated to be 3.8% and a lifetime risk of symptomatic knee OA in Western populations estimated to be over 40%.^{2,3} Knee OA can affect all three compartments of the knee in isolation or present in a multicompartmental pattern.

Osteoarthritis of the hip may also develop during early adulthood, with an age-standardised prevalence of 4.2% for symptomatic OA and 19.6% for radiographic OA.⁴ It is usually associated with adverse biomechanics arising from the developmental conditions hip dysplasia or femoro-acetabular impingement (FAI).^{5,6} Hip dysplasia describes a shallow acetabulum, whereas FAI results from subtle shape abnormalities of the femoral head or acetabulum.^{7,8} These conditions lead to abnormal pressures upon the joint cartilage that can progress to OA.^{7,8}

Osteoarthritis disease is a joint wide condition characterised by damage to the articular surfaces of the joint, bone sclerosis and cysts, stiffness of the soft tissues around the knee, synovitis and intermittent joint effusions. These features may be described as structural OA. Many patients with structural hip and knee OA have no or very mild symptoms from their joint. However, there is a group of patients with early or late OA structural changes who develop more pronounced hip or knee pain and loss of function, described as clinical OA, which includes a broad set of symptoms, characterised by pain and stiffness in the joint that produces a reduction in function and ultimately in quality of life (QoL).^{9,10} Genetic, acquired and environmental factors all play a role in the pathogenesis of the hip and knee OA.¹¹ In most cases, the natural history of the disease process is usually long and protracted, occurring over decades.

Current practice (National Institute for Health and Care Excellence osteoarthritis guideline)

All patients who present with OA should undergo initial treatment following the National Institute for Health and Care Excellence (NICE) clinical guideline CG226, with good education on their condition and treatment options; therapeutic exercise; and weight loss if appropriate. Other additional measures can be introduced to treatment such as topical or oral non-steroidal anti-inflammatory drugs (NSAIDs) (knee), intra-articular (IA) injections, manual therapy, and brace devices in some situations.¹² For patients with severe symptoms and end-stage (late) structural hip or knee OA, non-operative measures may not control symptoms sufficiently, and patients will discuss joint replacement (arthroplasty) as a possible treatment option.¹² The average age of these patients is mid to late 60s and joint replacement is successful for the majority of patients.^{13–15}

Health Technology Assessment project

This project investigates joint-preserving and joint replacement surgical options for patients with early-stage structural OA of the knee or hip. We start from the position that the NICE guideline treatments have all been tried before surgery is considered.

Patients with early OA of the knee or hip can also have marked symptoms that do not settle with non-operative treatment and continue to cause disability.¹³ We have previously shown that early knee OA patients can present with symptom levels as severe as those with end-stage disease who are considering joint replacement.⁹ These patients are typically a younger group of patients and are said to have entered the ‘treatment gap’, where symptoms and disability can be significant and yet treatment options are not so well established or widely used.¹⁶

Arthroplasty is a very effective treatment for pain but involves a major operation with a risk of complications. In younger patients, surgeons are reluctant to offer operations that will need to be revised in the future, typically after 10–15 years,

which is a complex operation with greater morbidity.¹⁷ The results are less reliable when structural knee OA is less severe, without exposed bone, even if pain is severe.^{18,19} There is uncertainty about whether arthroplasty should be used in early-stage structural OA or at younger ages, leading many patients to persist with pain and disability without a solution.²⁰

For many years, arthroscopic 'wash-out' of the knee joint was considered a first-line treatment in the early OA group. However, joint washout,²¹ together with meniscectomy,²² has been shown to be ineffective if indiscriminately used in the OA patient and the practice has greatly reduced.²¹⁻²³ Other arthroscopic procedures have become more widely used such as chondroplasty in the knee and osteochondral surgery for FAI in the hip.²⁴⁻²⁷ In these procedures unstable or abnormal cartilage and bone is removed. There are other procedures in clinical use but with variable uptake around the country. Osteotomies (realignment of the bones around the joint) can be used in the knee or hip, but questions remain about their ideal indication and benefit over a strategy of early arthroplasty.^{28,29} Also, procedures have been developed to stimulate new cartilage growth in the joint, or prevent further deterioration [such as chondroplasty, microfracture (MF), autologous chondrocyte implantation (ACI), or joint distraction].³⁰⁻³² Replacement of damaged areas can be done using allografts combining cartilage and a layer of subchondral bone.³³ There are also several forms of 'patch replacement' for small, localised defects with small synthetic resurfacing implants such as Episealer® (Episurf Medical, Stockholm, Sweden). However, for many procedures there is limited evidence, with variable descriptions of their use in early OA and no treatment guidelines specifically for this population. This has led to variations in commissioning and surgeon uptake, with uncertainty about the best approach to this common problem.²⁶

Given the apparent lack of evidence in this area of clinical care, the Health Technology Assessment (HTA) programme commissioned this study with the following aims and key objectives.

Aims of the project

The aim of this work is to establish if there is a role for surgery in the management of early structural OA of the hip or knee.

Key objectives

1. To undertake evidence synthesis to determine the clinical effectiveness of currently, but variably, used surgical interventions for treating or preventing early structural hip and knee OA, establishing what preoperative factors have an effect on the outcome or patient experience.
2. Where evidence of efficacy does exist to perform a health economic analysis to determine if surgical procedures are cost-effective.

Definition of early osteoarthritis of the knee and hip

A Cochrane review on the use of surgery in early knee OA has recently been completed. It helped formulate this application.³⁴ It had become clear through that process, which limits inclusions to randomised controlled trials (RCTs), that a wider search of study designs would be required to include observational studies, as recognised in the HTA brief for this study.

In the Cochrane review we spent considerable time defining 'early OA' and have extended this thinking to the early hip OA population for this HTA. OA is a joint wide process but in terms of stratifying surgical treatment pathways the state of the articular cartilage within the joint is central to choosing surgical treatment.³⁵ Therefore, we have used a broad definition with focus on the presence of pain and the exclusion of people with widespread full-thickness disease who would normally be considered suitable for arthroplasty surgery (most joint replacements are performed for clinical end-stage hip or knee OA, with extensive failure of joint structures such as major cartilage loss, bony eburnation and in

some cases bone loss). Early clinical hip or knee OA (EHKOA) represents the presence of symptoms and more focused structural change usually starting in one geographically distinct region of the joint. In addition, within the joint structural articular cartilage damage may be seen in combination with morphological features that predispose to early joint failure, such as dysplasia and femoro-acetabular impingement (FAI) in the hip and metaphyseal varus in the knee.^{7,8,36,37}

The Cochrane review on surgery for early knee OA used the following criteria for diagnosing OA on clinical grounds: knee pain for most days of the prior month, in addition to three of the following:

- Crepitus on active joint motion
- Morning stiffness > 30 minutes’ duration
- Age > 38 years
- Bony enlargement of the knee
- Bony tenderness of the knee on examination
- No palpable warmth

These diagnostic criteria for early OA reflect the currently used indications for surgical treatment. There is a very broad range of patients with reproducible patterns of disease both in hip and knee, where articular cartilage damage may be very minor through to focal full-thickness cartilage loss (FTCL), together with other joint changes. This cohort does not include individuals with end-stage OA joint failure where there is widespread cartilage loss and potentially eburnated bone or bone loss.

In this work, we focus on the need for a surgical ‘target’, defining early OA as (1) the presence of pain and (2) radiological evidence of structural early OA and (3) the absence of widespread disease. Diagnostic criteria for early structural hip and knee OA used in this study are seen in [Table 1](#) and are provided in more detail in the [Methods](#) section below.

We included interventions aimed at treating painful hip dysplasia and femoro-acetabular impingement (FAI) of the hip. Dysplasia of the hip and FAI are believed to represent an initial phase of the generation of hip OA.^{5,6} FAI describes subtle shape abnormalities of the hip where the femoral head is aspherical (cam morphology) or the acetabulum is deep (pincer morphology). The result is abnormal contact between the ball and socket, leading to joint damage and painful symptoms. Hip dysplasia describes a shallow hip acetabulum and can lead to early hip OA due to the resultant small contact area for load transmission between the femur and acetabulum. Therefore, these diagnoses were included in search for surgery used to treat early hip OA.

TABLE 1 Outline diagnostic criteria for early structural OA of the hip and knee

Hip: early structural OA, one or more of:
(1) KL grade 1–3 of weight-bearing anteroposterior (AP) radiograph
(2) PTCL or delamination on MRI.
(3) Focal FTCL in the hip on MRI
(4) Cartilage changes affecting the joint seen at arthroscopy or open surgery
Knee: early structural OA, one or more of:
(1) KL grade 1–3 of weight-bearing AP radiograph
(2) PTCL on MRI
(3) Focal full-thickness loss in one compartment of the knee on MRI > 2 cm ²
(4) Cartilage changes affecting the joint seen at arthroscopy or open surgery
KL, Kellgren-Lawrence grading system; ³⁸ MRI, magnetic resonance imaging; PTCL, partial-thickness cartilage loss.

Surgical treatment for early hip and knee osteoarthritis

As discussed above, surgical procedures for OA of the hip or knee are only indicated in patients who have clinical OA where both symptoms and structural OA change are present. This is most obviously seen in end-stage disease where the aim of joint replacement is to reduce symptoms. However, the same indications have been used in current surgical practice to treat early hip or knee OA. Here, the aims of the intervention are twofold: symptom relief and slowing structural progression to end-stage disease and potentially avoiding joint replacement surgery. Our understanding of the range of procedures that have been investigated both in practice and in research studies to treat early clinical hip and knee OA is derived from various sources, including our Cochrane review entitled *Surgical interventions for early structural knee osteoarthritis*, very recent scoping reviews (including RCT and cohort studies) of surgical procedures for treating early hip and knee OA performed in the preparation of this grant proposal and the supporting evidence and guidance provided within the HTA call.³⁴

As outlined in the HTA call there are two major surgical treatment types: joint preservation and joint replacement. Further to this, our preliminary work has identified four groups of intervention which encompass current procedures:

Joint preservation: non-regenerative procedures

In these procedures, the goal is to remove or repair pathological tissue within the joint, without aiming to regenerate new tissue. These procedures are aimed at reducing symptoms and reducing risk of progression.

In the knee (arthroscopic or open)

Procedures for early knee OA involve trimming (debriding) loose or damaged articular cartilage (chondroplasty) to smooth the surface using punches, shavers or radiofrequency ablation.²⁴ In the knee, meniscal damage is often associated with early knee OA. As a result, meniscal surgery aimed at treating the torn meniscus is sometimes performed in patients with early knee OA. Therefore, we have included the arthroscopic meniscectomy a procedure, but studies in people no detectable structural OA change.³⁹

In the hip: (arthroscopic or open)

As with the knee, procedures for early hip OA include debriding damaged articular cartilage, and either debriding or repairing damaged labrum. Where there is FAI, osteochondroplasty can be performed to reshape the hip and correct cam or pincer morphology at the same time as cartilage and labral procedures.⁴⁰ Where FAI has caused the articular cartilage to detach from the underlying bone, this can be secured back in position using fibrin glue. Osteochondroplasty can be performed with open surgery, but the vast majority of procedures are performed arthroscopically due to shorter recovery times and fewer complications.⁴¹ In this project, the term 'osteochondroplasty' is used to describe arthroscopic procedures of the knee or hip where bone or cartilage is debrided, or meniscus/labrum is debrided or repaired. The Cochrane review of FAI from 2014 by Wall *et al.* is currently being updated.

Joint preservation: biological reconstruction of the joint

Within surgical treatment of early hip and knee OA a number of techniques are used within both joints with the aim of replacing biological tissue within the joint. These principally consist of (1) subchondral bone marrow stimulation, (2) autologous cartilage reconstruction, (3) allografts of bone, cartilage or meniscus/labrum and (4) application of stem cells.

In the hip and knee all the following procedures can be adopted

Subchondral bone marrow stimulation encompasses a variety of techniques. The most widely adopted is MF, where the subchondral bone is perforated within a cartilage deficient region to release bone marrow and promote the formation of substitute cartilage (fibrocartilage). This may be supplemented by a synthetic matrix [autologous matrix-induced chondrogenesis (AMIC)].⁴²

Autologous chondrocyte implantation, now mainly third generation matrix-associated autologous chondrocyte implantation (MACI), involves the application of cultured chondrocytes (cartilage producing cells) to regenerate cartilage in areas where it is deficient. While this was appraised recently by NICE, the scope limited the assessment to licensed products that have now been overtaken by other technologies on the market, and other technologies are emerging in the literature.⁴³ It was also focused on isolated chondral lesions rather than OA, a wider topic.^{43,44} Uncertainty remains

about the cost-effectiveness of ACI in early OA and the criteria for determining if OA is too advanced for the procedure to work.⁴³ The Warwick assessment report for NICE concluded that ACI was not cost-effective in early OA at current prices, but might be if the cost was reduced.

Meniscal/labral and osteochondral allografts (OCAs) can be employed. Meniscal allografts in the fully or functionally menisectomised patient are thought to reduce later OA and improve symptoms.⁴⁵ Studies also report reconstruction of the hip labrum using allograft.⁴⁶ OCAs replace both chondral cartilage and the underlying subchondral bone, which is often abnormal. Their use is increasing in the UK due to emerging long-term follow-up data and observational studies, largely from early adopters in the USA.⁴⁷

Stem cells (usually harvested from liposuction or bone marrow) have been used in a few studies but have yet to be properly appraised.⁴⁸ They were not included in the NICE appraisal of ACI.

Matrix implants have been used with or without stem cell treatments. In these treatments, man-made synthetic tissue is implanted to reconstruct the joint.

Knee joint distraction (KJD) is another technique that temporarily separates the tibia and femur, unloading the knee joint and stimulating reconstruction of the cartilage thickness.

Scaffolds – several multilayered ‘scaffolds’ have been used for reconstruction of chondral or osteochondral defects (e.g. MaioRegen, Agili-C®). These will also be included.

Joint preservation: load-modifying procedures

Osteoarthritis is frequently associated with adverse joint loading, in particular coronal malalignment of the knee and dysplasia of the hip. Load-modifying procedures in both joints are used to optimise biomechanics of reducing symptoms and potentially delaying progression.

In the knee

An osteotomy used to treat knee OA is a realignment procedure aimed at transferring the weight-bearing region of the knee joint from the diseased region to a disease-free region. This is achieved by cutting and reshaping the bone in the proximal tibia [high-tibial osteotomy (HTO)] or distal femur [distal femoral osteotomy (DFO)]. A previous Cochrane review has highlighted their potential.²⁸ Chronic medial compartment joint loading can be modified using an extracapsular device. Devices such as the KineSpring® (Moximed, Inc, Hayward, CA, USA) have been developed to partially absorb the load passing through the medial compartment without compromising the joint surface.⁴⁹ Metallic interpositional devices are designed to restore medial joint space narrowing (JSN) caused by OA. In so doing, a varus alignment can be reduced.⁵⁰ They can be inserted using minimally invasive techniques and have the advantage of being inserted without the need for resection of underlying bone.

In the hip

For dysplasia, peri-acetabular osteotomies (PAOs) can be performed to reorientate the acetabulum and create a larger contact area for load transmission.²⁹ When a PAO is not indicated, an alternative is to rotate bone from above the socket to increase the depth of the acetabulum (shelf acetabuloplasty).⁵¹ Some patients develop hip pathology secondary to abnormal rotation or alignment of the femur that can also be corrected with osteotomies. Osteotomies of the pelvis or femur represent major surgery with a prolonged period of reduced mobility postoperatively.

Joint replacement procedures

Total hip or knee replacement (THR or TKR) involves replacing the surfaces of the knee joint with artificial implants. Unicompartmental knee replacement (UKR) only involves replacing the surfaces of a single diseased compartment (medial, lateral or patellofemoral) rather than the whole joint surface. Resurfacing of the hip involves placing a new metal surface over the head of the hip and not implanting a stemmed component. A HTA report on metal-on-metal hip replacements also covered resurfacing, based on evidence published before the end of 2012. This underpinned NICE Technology Appraisal guidance 304 issued in February 2014.

Arthroplasty (joint replacement) surgery is not routinely offered to those with early structural hip or knee OA because there is some evidence that it is associated with a poorer outcome when end-stage bone-on-bone changes are not present.^{18,19} Arthroplasty prostheses may need to be replaced after 10–15 years, and the revision procedure is more difficult with more postoperative morbidity, so surgeons are reluctant to carry out TKR or THR in people under 50 years. Many patients develop early structural OA much earlier than this, for example, because of knee injuries or hip dysplasia. We will examine the evidence for joint replacement surgery in those with early structural hip or knee OA.

Several metal or plastic 'patch' replacements [also known as mini-implants], e.g. Episealer, BioPoly® (BioPoly, Fort Wayne, IN, USA) have been used in the knee and hip to treat cartilage damage and early OA so this intervention is included in the study.

Chapter 2 Evidence synthesis methods

There has been patient and public involvement (PPI) in this study in the original application to National Institute for Health and Care Research (NIHR) and in the process of delivering the study. There was a specific PPI representative in the study team involved in the planning and delivery of the work.

Evidence synthesis protocol

This systematic review protocol was reported in accordance with the Preferred Reporting Items for Systematic review and Meta-Analysis Protocols (PRISMA-P).⁵² The review has been reported in accordance with the PRISMA guidelines.⁵³ The protocol was registered with the International Prospective Register of Systematic Reviews (PROSPERO; CRD 42019156444).

Criteria for inclusion and exclusion of studies

Types of studies

From the scoping review and as anticipated in the Commissioning Brief, we expected that there were few RCTs and note that much research is conducted as case series, often reported and summarised in systematic reviews.

We included high-quality, recent systematic reviews, defining 'recent' as the last 5 years. We planned that if we did not find at least one high-quality review, relevant to our scope, published since 2014, we would search back till 2010.

We included RCTs and quasi-RCTs, longitudinal non-randomised studies (NRS) with a comparison of interventions (e.g. cohort studies, case-control studies, controlled before-and-after studies, controlled interrupted time series and observational studies with a reasonably matched group not having surgery from inception of the bibliographic databases). Quasi RCT were defined as a trial where participants were allocated to different treatment groups using a non-random method. We also identified observational non-comparative studies from systematic reviews.

We classified included studies into three hierarchical groups:

1. RCTs/quasi-randomised controlled trials (q-RCTs) together with systematic reviews of these study designs or mixed designs
2. comparative NRS and systematic reviews of these designs
3. systematic reviews of non-comparative study designs (e.g. case series) or individual case series in the absence of systematic reviews.

We excluded cross-sectional studies that simply report current states in people that have had past interventions, with no baseline data. We excluded case reports. Observational studies were considered important for longer-term outcomes (5 years or more), including the need for further surgery, including arthroplasty and revision arthroplasty and adverse event (AE) outcomes (e.g. infections, thromboembolic disease, removal of grafts, 30-day mortality).

We included studies published solely as conference abstracts, with caveats, as appropriate, from 2017 onwards and contacted authors for further information, as appropriate. Unpublished data were included only if authors give permission for publication. There was no language restriction but only studies in English were primarily included. The purpose of having no language restriction was to check what proportion of studies which were published in English. Many non-English language studies had English abstracts. However, if it appeared that important evidence on a significant technology is not available in English, we used translation. Registry data were used for trends in arthroplasty by age, and longer-term outcomes.

We did not routinely data extract studies with fewer than 20 patients but considered reported them narratively if the intervention was considered likely to become important, for example, pilots of new technologies requiring future research. We looked for studies with a minimum follow-up period, decided for each outcome, but will not be < 6 months. For example, relief of symptoms may be shown by 6-month data, but radiological progression required at least 2–3 years' follow-up.

Setting

The setting was secondary care for surgical procedures and primary or secondary care for non-operative comparators.

Target population

We included studies in adults (18 years of age and older) with a diagnosis of early structural OA of the hip or knee joint [early knee osteoarthritis (EKOA)/early hip osteoarthritis (EHOA)], or conditions very likely to progress to OA, namely FAI, and dysplasia in the hip, and articular cartilage injury. We included studies in people with pain and no detectable structural radiographic change if they also have morphological features or tissue damage, which often leads on to OA.

We defined morphological features as features that are believed to be predisposing factors for the development of arthritis: in the hip, femoro-acetabular impingement (FAI) (including pincer, cam and combined bony features) and acetabular dysplasia; in the knee, tibia vara and varus alignment.

We excluded studies in people with pain but no detectable structural change and no morphological features.

We defined EKOA and EHOA as:

The presence of joint pain and one of the following features of structural disease:^{35,54,55}

(a) Radiographic

Knee:

Kellgren-Lawrence (KL) grade 1, 2, 3 or equivalent. A grade of 3 was consistent with early OA depending on the precise definition used by the authors, which is not consistent across the literature. The KL grade has only moderate correlation to structural severity, and the boundary between 2 and 3 is often particularly difficult to define as the interpretation of 'possible' and 'definite' JSN can be very subjective and 'definite' narrowing may still represent early OA in some studies.^{56–60} We therefore also planned to include those with the KL3 of the knee in the absence of exclusion criteria outlined below.

Hip

Kellgren-Lawrence grade 1, 2, 3 or equivalent (e.g. Tönnis), following the same principle as above.⁶¹ For hip dysplasia: centre-to-edge angle CEA < 20° (mild 15–20°, moderate 10–15° and severe < 10°) and for FAI: Alpha angle > 55° or centre-edge angle of > 40°.^{8,62}

Kellgren (1963) described 4 grades of hip OA, for which 1, 2 and 3 are:

- Grade 1: (doubtful OA), possible narrowing of the joint space medially and possible osteophytes around femoral head
- Grade 2: (mild OA), definite narrowing of the joint space inferiorly, definite osteophytes and slight sclerosis
- Grade 3 (moderate OA), marked narrowing of the joint space, slight osteophytes, some sclerosis and cyst formation, and deformity of the femoral head and acetabulum.

We note that there was significant variation in how researchers interpret/modify the KL grade and that it is often simplified to:

1. Possible osteophytes with no joint space width (JSW) narrowing.
2. Definite osteophytes with possible JSW narrowing.

3. Definite osteophytes and marked JSW narrowing, some sclerosis and cyst formation.

(b) *MRI* (magnetic resonance imaging)

Cartilage degradation but minimal full-thickness loss (< 10% of surface area or < 1 cm).

For example, **Boston Leeds Osteoarthritis Knee Score (cartilage morphology)**.⁶³

- Any level of cartilage loss (per cent of subregion surface area affected by cartilage loss) (grade 1: < 10%; grade 2: 10–75%; grade 3: > 75%).
- And minimal full-thickness loss (full-thickness loss score): grade 0: none; grade 1: < 10%.

Excluded are studies in people with full-thickness loss $\geq$ 10%

OR

Whole organ magnetic resonance imaging score cartilage morphology.⁶⁴

- Grade 1: increased signal from normal-thickness cartilage
- Grade 2: partial-thickness focal defect < 1 cm
- Grade 2.5: full-thickness focal defect < 1 cm
- Grade 3: multiple areas of partial-thickness defect within areas of normal thickness
- Grade 4: diffuse partial-thickness loss

Excluded are studies in people with:

- Grade 0: normal thickness and signal
- Grade 5: multiple areas of full-thickness loss (grade 2.5) or a grade 2.5 lesion wider than 1 cm but < 75% of the region
- Grade 6: diffuse ($\geq$ 75% of the region) full-thickness loss

OR

Magnetic resonance imaging osteoarthritis knee score cartilage morphology or magnetic resonance imaging osteoarthritis hip score.⁶⁵

- Any cartilage loss as a % of surface area (grade 1: < 10%, grade 2: 10–75%, grade 3: > 75%).
- And extent of FTCL in a region: (grade 0: < 10%, grade 1: < 10%)

Excluded are studies in people with FTCL $\geq$ 10%.

OR any equivalent MRI scoring system.

(c) *Arthroscopic*

A single area of full-thickness cartilage damage assessed with the International Cartilage Repair Society (ICRS) score:⁶⁶

- Grade 1: soft indentation (A – patient assessed) and/or superficial fissures and cracks (B – surgeon assessed)
- Grade 2: lesions extending down to < 50% of cartilage depth
- Grade 3: cartilage defects extending down > 50% of cartilage depth (A) as well as down to calcified layer (B)

(Reproduced from the ICRS Score).⁶⁶

Excluded are studies in people with grade 4 ICRS: through the subchondral bone.

OR

An equivalent arthroscopic scoring system such as Outerbridge scale.⁶⁷

- Grade 1: cartilage with softening or swelling
- Grade 2: partial-thickness defect with fissures on the surface that do not reach subchondral bone of exceed 1.5 cm in diameter
- Grade 3: fissuring to the level of subchondral bone in an area with a diameter > 1.5 cm

Excluded are studies in people with grade 4 Outerbridge: exposed subchondral bone.

Exclusion criteria

We excluded studies in participants with the following characteristics:

- Asymptomatic individuals.
- End-stage OA with FTCL (> 2 cm² area or ≥ 10%), bony deformity or bone-on-bone change on radiographs (KL grade 4).
- Full-thickness loss of cartilage in more than one compartment of the joint (knee).
- History of trauma and/or surgery to the knee/hip for fracture.
- The diagnosis of an inflammatory arthritis, metabolic bone disease.

Where research has combined patients with early and late hip or knee OA, we included outcomes that related to the early disease state, if available. We excluded studies with < 80% of participants with early OA unless the results for people with early OA are reported separately.

Interventions

It was anticipated that there would be instances where concomitant procedures in addition to the joint preservation or surgical intervention were reported.

Joint preservation: non-regenerative procedures

1. Knee.

We included the following procedures:

- Arthroscopic debridement, trimming and surface smoothing
- Chondral debridement
- Chondroplasty
- Meniscal debridement
- Meniscal reconstruction
- Meniscal repair/meniscal surgery
- Meniscectomy
- Meniscal implants
- Root repair

2. Hip

We included the following procedures and additional terms:

- Cam debridement
- Head neck junction reshaping

- Labral debridement
- Labral grafting/reconstruction Labral repair in people with FAI or dysplasia (but not labral tears in isolation)
- Osteochondroplasty
- Recontour
- Debridement, trimming, surface smoothing

Joint preservation: biological reconstruction

1. Knee

We included the following procedures:

- Allogenic chondrocytes
- Allografts (bone, cartilage, meniscus)
- Autologous cartilage reconstruction
- Autologous chondrocyte implantation (ACI)
- AMIC
- Bone marrow concentrate (BMC)
- Chondral grafts and allografts
- Implant
- Distraction surgery
- KJD
- MACI
- Meniscal allografts
- MF
- Mosaicplasty
- Multilayer scaffolds
- Subchondral injections of bone substitute
- Chondral fleece, patch or scaffold
- Scaffolds

These techniques can be recorded under the following trade/brand names: Agili-C, CartiONE™ (Cartilage Repair Systems B.V., Driebergen-Rijsenburg, The Netherlands), ChondroCelect (TiGenix N.V., Leuven, Belgium), Chondrosphere/Spherex (ReLive Biotechnologies, Newton, Massachusetts, USA), MACI (Vericel Corporation, Cambridge, Massachusetts, USA), MaioRegen (Fincera Faenza S.p.A., Faenza, Italy), Novocart (TETEC AG, Reutlingen, Germany), OsCell [OxCell Ltd., Manchester, Greater Manchester, UK (now dissolved)], Syngenit (Synergy Medical Technologies Ltd, Brighton, UK).

2. Hip

We included the following procedures:

- Allogenic chondrocytes
- Allografts (bone, cartilage, labrum)
- Autologous chondrocyte implantation (ACI)
- AMIC
- BMC
- Labral allograft
- Matrix implant
- MACI
- MF
- Mosaicplasty
- Osteochondral grafts and allografts
- Stem cell injection, provided targeted at a specific site/lesion

- Subchondral bone marrow stimulation
- Osteochondral fleece, patch or scaffold.
- Scaffold

These techniques can be recorded under the following trade/brand names; Agili-C, CartiONE, ChondroCelect, Chondrosphere/Spherox, MACI, MaioRegen, Novocart, OsCell and Syngenit.

Joint preservation: load-modifying procedures

1. Knee

We included the following procedures:

- DFO
- Extracapsular device
- HTO
- Extracapsular device
- Metallic interpositional devices
- Osteotomy
- Realignment

These techniques could be recorded under the following trade/brand names; KineSpring (MISHA Knee System, Moximed Inc, Hayward, California, USA), Atlas [Moximed] (AtlasMoximed Inc, Hayward, California, USA), Tomofix (DePuy Synthes, aynham, Massachusetts, USA), and Optyline (NuVasive Specialized Orthopedics, San Diego, California, USA).

2. Hip

We included the following procedures

- Acetabular osteotomy
- Cheilectomy
- Derotation osteotomy
- Ganz osteotomy
- Ganz trochanteric flip
- Osteoplasty
- Osteotomy
- PAO
- Shelf acetabuloplasty
- Triple osteotomy
- Upper femoral osteotomy
- Valgus osteotomy
- Varus osteotomy

Joint replacement procedures

1. Knee

We included the following joint replacement and mini-implant procedures:

- Arthroplasty prostheses
- Partial knee replacement
- Patellofemoral joint replacement
- UKA/UKR (medial, lateral or patellofemoral)

- TKR/total knee arthroplasty (TKA)
- Metal patches
- Mini-implants
- Patch replacements

These techniques can be recorded under the following trade/brand names: BioPoly, Episealer Episurf, Hemicap (HemiCAPArthroSurface, Inc., Franklin, Massachusetts, USA) and Unicap (ArthroSurface, Inc., Franklin, Massachusetts, USA).

2. Hip

We included the following procedures:

- Arthroplasty prostheses
- Metal-on-metal hip replacement
- Hip prosthesis
- Resurfacing of the hip
- Total hip replacement (THR)/total hip arthroplasty (THA)
- Metal patch
- Mini-implants
- Patch replacements
- Small surface replacement

These techniques can be recorded under the following trade/brand names: BioPoly, Episealer, Episurf, Hemicap and Unicap.

We excluded studies of prostheses that are no longer used, and studies of revisions of surgical procedures (as interventions).

Comparator interventions

Our starting position was that conservative management is as defined by the NICE Clinical Guideline on osteoarthritis (CG226, 2022), which is similar to what is in the Commissioning Brief.¹² We note the NICE guideline was been updated (publication in 2022) and that some of the recommendations may be changed in the light of new evidence (e.g. the non-use of electrotherapy and IA injection). Although we are aware that other OA treatment guidelines are available starting from CG226 is justified since all individuals with hip and knee OA in the UK should be offered NICE recommended interventions, and many of the surgical interventions examined in this review will have been compared to a non-operative comparator, or would have to be if we are to make research recommendations. We assumed that the following have been offered before surgery is considered:

- Education and information
- Activity and exercise
- Weight loss interventions, in indicated
- Analgesics (paracetamol and topical NSAIDs)

Accordingly, studies were included if the comparators consist of one or more of the following 'adjunct' interventions:

- Physiotherapy, including therapeutic exercise.
- Pharmacological management (e.g. NSAIDs), IA injection therapy (e.g. steroids).
- Bracing, orthosis and assistive devices.

Studies were eligible where these non-operative interventions (defined by NICE as adjuncts) are delivered as comparators to surgical interventions. We also included comparisons with no treatment or NICE 'core' treatments (education, exercise, weight loss and analgesia). We did not include interventions that the NICE guideline recommended against using (and that are not being updated) – that is, acupuncture, electrotherapy, topical rubifacients (such as trolamine salicylate) and glucosamine or chondroitin.

Studies comparing one class of surgery (non-regenerative procedures, biological procedures, load-modifying procedures and joint replacement) were included as long as one form of surgery has been compared with non-surgical care, directly or indirectly:

- Joint-preserving surgical intervention (non-regenerative procedures, biological reconstruction procedures, load-modifying procedures) versus joint replacement.
- Non-regenerative surgical procedures versus biological reconstruction surgery procedures.
- Non-regenerative surgical procedures versus load-modifying surgical procedures.
- Biological reconstruction surgery procedures versus load-modifying surgical procedures.

We also included studies which compare one type of surgical procedure with another type of surgical procedure within the same class of surgical intervention.

We included NRS where there is comparison with a non-operative treatment approach, using the criteria outlined above. We did not include NRS where the comparison of surgery was with another surgical intervention, either in the same class or in a different class of intervention.

Outcomes

We included outcome measures which satisfied the Outcome Measures in Rheumatoid Arthritis Clinical Trials (OMERACT) and Osteoarthritis Research Society International (OARSI) core outcome set for hip and knee OA (and the OMERACT Core Outcome Domain Set for joint replacement).⁶⁸⁻⁷¹ These included clinical outcomes, radiological outcomes and resource use.

We included both patient-reported outcome measures (PROMs) and clinician-reported outcome measures. We extracted data from validated PROMs for people with hip and knee OA or cartilage disease, including: Western Ontario and McMaster Universities Arthritis Index (WOMAC) (total score); Oxford Knee Score (OKS); Oxford Hip Score (OHS), Hip Disability and Osteoarthritis Outcome Score (HOOS), Knee Injury and Osteoarthritis Outcome Score (KOOS); any other validated PROM, including those for non-arthritic joint symptoms [e.g. Non-Arthritic Hip Score, International Hip Outcome Tool (iHOT), Hip Outcome Score (HOS), International Knee Documentation Committee (IKDC)]. We used minimum clinically important differences (MCIDs) when interpreting changes. Clinician-reported outcome measures included the Harris Hip Score (HHS), American Knee Society Score (AKSS), Hospital for Special Surgery (HSS) Score.

Combined pain, physical function and other symptoms scores. If data on more than one scale were provided, we extracted data, giving preference to the first four measures.

1. WOMAC total score
2. OKS or OHS
3. HOOS or KOOS. Also HOOS-JR and KOOS-JR
4. HHS
5. HSS Knee Score
6. HOS
7. IKDC questionnaire
8. Non-Arthritic Hip Score
9. iHOT

10. AKSS
11. Copenhagen Hip And Groin Outcome Score (HAGOS)

Pain measured using a validated pain score or subscale of other measure. If data on more than one pain scale were provided for a trial, data were extracted according to hierarchy presented below.^{72,73}

12. Global pain (pain overall, including at rest, such as interrupting sleep)
13. Pain on walking
14. WOMAC pain subscale
15. Composite pain scores other than WOMAC
16. Pain on activities other than walking
17. Rest pain or pain during the night
18. WOMAC global algofunctional score
19. Lequesne OA index global score
20. Other algofunctional scale
21. Patient's global assessment
22. Physician's global assessment
23. Other outcome
24. No continuous outcome reported

Physical function measured with a validated physical function tool or subscale of other measure. If data on more than one physical function scale were provided for a trial, data were extracted according to hierarchy presented below:⁷²

25. Global disability score
26. Walking disability
27. WOMAC functional limitation subscale
28. Function in daily living subscale for KOOS or HOOS
29. Disability other than walking
30. WOMAC global scale
31. Lequesne OA index global score
32. Lysholm or Lysholm/Tegner
33. IKDC
34. Return to sport and level achieved compared to pre-injury
35. Return to other activities such as occupational
36. Other functional scales
37. Patient's global assessment
38. Physician's global assessment

Surgical complications and AEs, for example, infections, thromboembolic disease (deep-vein thrombosis/pulmonary embolism), removal of grafts (for some biologics), 30-day mortality.

Patient-reported experience measures (PREMs) measured using any validated PREM, for example, CARE (Consultation And Relational Empathy) Patient Feedback measure.

Health-related quality of life (HRQoL) measured using a validated score such as the EuroQol-5 Dimensions (EQ-5D), Short Form questionnaire-12 items (SF-12); or any other validated HRQoL measure.

Joint replacement, representing failure of initial surgical procedure (progression of OA 2 years' follow-up or more)

Reoperation rate or revision of initial surgical procedure, including any new procedure, and time to revision or other reoperation.

Radiological outcomes (progression of OA – 2 years' follow-up or more) Radiographic measures: Minimum JSW or validated alternative (e.g. mean JSW, joint space area).⁷⁴

MRI measures. Morphological: cartilage volume, cartilage thickness, surface area. Compositional: T2 mapping, T1 Rho, delayed Gadolinium-Enhanced MRI of Cartilage (dGEMRIC).

Semi-quantitative measurements, including KL grade or MRI scores.⁷⁵

Resource use to inform the health economic modelling. The following use was assessed: operations, non-operative treatments, inpatient hospital stays, outpatient visits, diagnostic imaging and tests, prescriptions, community care consultations, indirect costs – for example, loss of productivity was explored by extracting work-related outcomes such as loss employment, return to work and any other employment-related outcomes.

Minimum duration of follow-up

Due to the multi-dimensional outcomes we assessed (e.g. early complication, symptom improvement, further surgery, survival of implants), there was no single minimum duration of follow-up, but we excluded studies with < 6 months' follow-up. Where multiple time points were reported, we categorised them into short- (6 months to 2 years), intermediate- (more than 2 years to < 5 years), long-term (5 years or more) follow-up, and then analysed and reported separately as appropriate. We used the longest follow-up reported within each follow-up category, but it was anticipated that there may be situations where data from several time points were used to estimate areas under curves.

Search strategy

We searched the following sources:

- Full-text publications in MEDLINE, Excerpta Medica dataBASE (EMBASE), Cochrane Library back to inception
- Meeting abstracts back to 2017 in EMBASE
- HTA database www.crd.york.ac.uk/crdweb
- NICE www.nice.org.uk/
- International Network of Agencies for Health Technology Assessment (INAHTA) www.inahta.org/Publications/
- Scottish Intercollegiate Guidelines Network (SIGN) www.sign.ac.uk/
- Canadian Agency for Drugs and Technologies in Health (CADTH) www.cadth.ca/
- Institute for Quality and Efficiency in Health Care (IQWiG) www.iqwig.de/en/methods/results/hta-reports.7459.html
- Research in progress: ClinicalTrials.gov <https://clinicaltrials.gov/> and WHO Clinical Trials Search Portal <http://apps.who.int/trialsearch/> and NIHR awards <https://fundingawards.nihr.ac.uk/search>

We established monthly auto-alerts in MEDLINE for full text and EMBASE for meeting abstracts only. We checked studies included in systematic reviews found in scoping searches and main searches.

We searched all databases for RCTs and comparative studies from their inception to present day. We searched for systematic reviews, initially from 2014, and if there are no high-quality reviews, we searched back to 2010. We imposed no restriction on language of publication. All reference lists from those studies, which were judged as potentially eligible on screening, were reviewed.

For commercial technologies, we searched manufacturers' websites for unpublished research, contacting them for further information when appropriate.

Search strategies for Ovid MEDLINE are shown in [Appendix 1](#).

Data collection and analysis

Selection of studies

One person performed an initial screen of titles and abstracts of all potential studies identified as a result of the search to exclude clearly irrelevant records. Two reviewers then independently screened the remaining titles and abstracts and coded them as 'retrieve' (eligible or potentially eligible/unclear) or 'do not retrieve'. Disagreements were resolved by discussion, and if necessary, through involvement of a third reviewer. We retrieved the full-text study reports/publication, two reviewers then independently screened these full text and identified studies for inclusion, using an eligibility checklist, and identified and recorded reasons for exclusion of the ineligible studies. Clinical leads in Oxford and Warwick were asked to review inclusions and exclusions where there is uncertainty regarding eligibility.

We applied a hierarchy of evidence in deciding which studies to analyse: firstly sorting the studies into three categories according to:

1. RCTs/q-RCTs and systematic reviews of these study designs or mixed designs that include RCTs and q-RCTs.
2. Comparative NRS and systematic reviews of these designs.
3. Systematic reviews of other non-comparative study designs (i.e. case series). We excluded studies with fewer than 20 participants.

We synthesised the RCT/q-RCT evidence initially, using existing systematic reviews as appropriate, adding evidence from comparative NRS if there are insufficient RCTs. In most topics, our scoping searches showed that reviews were plentiful, but there were few RCTs or comparative studies. Where there were good-quality reviews (defined as meeting most of the quality criteria), we checked whether the included studies were in people with early OA, as defined by our protocol. If studies recruited people with a mixture of stages of OA, we treated them as excluded unless results for people with early OA were reported separately. We checked the primary studies we identify against those reported in the systematic reviews and added any new studies to the review. Then we decided whether to report the original review conclusions, or update the review, or conduct a de novo systematic review.

This evidence from (1) and (2) was used as the clinical effectiveness evidence, which was reported in summary of findings and Grading of Recommendations Assessment, Development and Evaluation (GRADE) tables. It was then used in the health economic modelling if it was of adequate quality. Where there was no evidence in (1) and (2), an absence of evidence will be reported in the tables.

We resolved any disagreements through discussion or by involving additional members of the team if necessary. Multiple reports of the same study were collated so that each study, rather than each report, is the unit of interest in the review.

Documenting exclusions

Study flow was recorded using a PRISMA flow diagram.

Data extraction and management

We used a standard data collection form for study characteristics and outcome data that had been piloted on at least one study in the review. Data were extracted by one reviewer and checked by a second. Any disagreements were resolved through discussion or by involving additional members of the team if necessary. We extracted the following study characteristics:

1. **Methods:** study design, total duration of study, details, number of study centres and location, study setting, withdrawals and date of study.
2. **Participants:** *N*, mean age, age range, sex, body mass index (BMI), severity of condition (i.e. structural disease category/type), diagnostic criteria (i.e. grade of structural change), prior surgery, previous treatments for OA (especially NICE core), limitation of normal activities such as sport, previous surgical procedures, inclusion criteria and exclusion criteria.

3. **Interventions:** total number of intervention groups within each study, specific details of each intervention, concomitant procedures (especially NICE core), and comparator, who provided and conducted the intervention, materials used, standardisation of the surgical and non-surgical interventions and adherence to the standardised technique, any modifications or tailoring of the intervention during the trial period. We considered the expertise of centres by looking for data on numbers having procedures per year and how long the procedure has been carried out.
4. **Outcomes:** outcomes from our list, whether specified and collected, and time points reported.
5. **Declarations:** funding for trial, and notable declarations of interest of trial authors.

We extracted the number of events and number of participants per treatment group for dichotomous outcomes, and means and standard deviations (SDs) and number of participants per treatment group for continuous outcomes. We resolved disagreements by consensus or by involving a third review author. For included systematic reviews of comparative studies, we double-checked that data were entered correctly by comparing the data presented in the systematic review with the original study reports. Where both final values and change from baseline values are reported for a given outcome, we extracted the baseline and final values.

Additional information was extracted to inform the health economic modelling, including type of model and justification, mapping methods used to derive health utilities, deterministic incremental cost-effectiveness ratios (ICERs), sensitivity and scenario analyses.

Assessment of risk of bias in included studies

One reviewer assessed risk of bias for each study, and a second checked the assessment.

The quality of systematic reviews was assessed using the NIH criteria;

1. Is the review based on a focused question that is adequately formulated and described?
2. Were eligibility criteria for included and excluded studies predefined and specified?
3. Did the literature search strategy use a comprehensive, systematic approach?
4. Were titles, abstracts and full-text articles dually and independently reviewed for inclusion and exclusion to minimise bias?
5. Was the quality of each included study rated using a standard method to appraise its internal validity?
6. Were the included studies listed along with important characteristics and results of each study?
7. Was publication bias assessed?
8. Was heterogeneity assessed? (this question applies only to meta-analyses)

Y, yes; N, no; CD, cannot determine; NA, not applicable; NR, not reported.

In addition, we reported whether authors considered clinical significance and not just statistical significance.

For RCTs and q-RCT, risk of bias was to be assessed using the Cochrane Risk of Bias tool version 2.0.⁷⁶ The tool covers five possible bias domains (bias arising from the randomisation process, bias due to deviations from intended interventions, bias due to missing outcome data, bias in the measurement of the outcome and bias in the selection of the reported result) which have been shown to reflect the main mechanisms for bias in RCTs.⁷⁶ The assessment of risk of bias was only evaluated for our main outcomes: (1) combined pain and physical function, (2) pain and (3) function. Bias due to missing outcome data were evaluated separately by category of outcome follow-up times (short, medium, long). Each domain was assigned a judgement of 'high', 'low' or 'some concerns'. Narrative descriptions were used to provide reasons for a particular judgement.

For risk of bias in comparative NRS, we used tools developed by the National Institute for Health, National Heart, Lung and Blood Institute (NIH NHLBI). These tools focus on concepts including biases (selection, performance, detection and attrition), confounding, power and strengths of associations between treatments and outcomes. There are different tools for each major type of study design, such as for cohort studies (two groups), before-and-after

studies (one group) and case-control studies (two group studies looking at associations). Specific guidance notes are provided for each tool. Each question was assigned a response of 'yes', 'no', 'can't determine', 'not reported' or 'not applicable'. The study was then assessed for overall quality (good, fair, poor) based on the responses to the individual questions, where a good study has the least risk of bias, and results are considered to be valid; a fair study is susceptible to some bias, but this is not deemed sufficient to invalidate the results; and a poor study indicates that the study is at a significant risk of bias (reproduced from <https://methods.cochrane.org/bias/resources/rob-2-revised-cochrane-risk-bias-tool-randomized-trials>).

In general, we avoided providing a single score of criteria met/not met, because some criteria are more important than others (especially confounding). So we used text such as good, fair, poor. However, as a rough guide to these terms, for the NIH checklist, we used the number of 'yes' responses as a general rule of thumb to judge the overall quality of a study as follows:

- For before-and-after studies (10 core questions): Good 8–10, Fair 5–7, Poor < 5.
- For cohort studies (14 questions), Good 10–14, Fair 7–9, Poor < 7.

Some studies excluded on quality grounds were reported narratively as additional analysis, if thought to have useful information such as identifying research needs.

Main planned comparisons

We planned to analyse results for the following main comparisons, separately for EKO and EHO and by type of surgical procedure:

1. Surgical intervention (non-regenerative procedures, biological reconstruction procedures, load-modifying procedures, joint replacement procedures) or combinations of surgical interventions versus non-operative management following NICE core guidance (i.e. combination of information, exercise/physical therapy and weight loss).
2. Surgical intervention versus non-operative management following a single NICE core therapy intervention (i.e. information or exercise/physical therapy or weight loss). We planned to check the amount of treatment received, for example, the number of physiotherapy sessions.
3. Surgical intervention versus non-operative management following NICE guidance (i.e. combination of information, exercise/physical therapy or weight loss) plus one or more adjunctive non-surgical component (such as oral NSAIDs or injection therapy).
4. Surgical intervention versus any pharmacological management (e.g. NSAIDs, injection therapy).
5. Surgical intervention versus bracing/orthosis or assistive devices.
6. Surgical intervention versus electrotherapy (e.g. transcutaneous electrical nerve stimulation, ultrasound).
7. Joint-preserving surgical intervention (non-regenerative procedures, biological reconstruction procedures, load-modifying procedures) versus joint replacement.
8. Non-regenerative surgical procedures versus biological reconstruction surgery procedures.
9. Non-regenerative surgical procedures versus load-modifying surgical procedures.
10. Biological reconstruction surgery procedures versus load-modifying surgical procedures.
11. Surgical intervention type 1 versus surgical intervention type 2 (within a surgical procedure group).

Data synthesis and measures of treatment effect

Details of included studies were combined in the review by hip or knee then according to the comparison between intervention and comparator, and then outcomes. We synthesised and reported the data separately for comparative and non-comparative studies; data from RCTs were also reported and analysed separately to non-randomised comparative studies.

We planned to undertake meta-analyses only where appropriate based on an assessment of clinical and methodological heterogeneity. Where studies were considered sufficiently homogeneous and data were available, meta-analysis was

performed using a random effects model. We analysed data for each outcome separately for EKO and EHO for each of the main comparisons and type of surgical procedure. Where considered appropriate, data on very similar surgical procedures, and or comparator, were combined in the same analysis.

Statistical heterogeneity was assessed by visual inspection of the forest plot for obvious differences in results between the studies, and by using the I^2 and chi-squared statistical tests. If we identified substantial statistical heterogeneity, we reported it and investigated possible causes, and unexplained heterogeneity was reflected in the GRADE assessment. We provided narrative summaries instead of meta-analysis results, if more appropriate to do so. Decisions on meta-analysis/narrative summary were also guided by visual inspection of forest plots, since heterogeneity can be in both direction and effect size. Pooled estimates of treatment effect were obtained using Cochrane RevMan (The Cochrane Collaboration, The Nordic Cochrane Centre, Copenhagen, Denmark) software (version 5.3). We presented data using forest plots where possible, even if pooling is not carried out.

We analysed continuous data (such as pain, function, HRQoL, PROMs and PREMS) as mean difference or standardised mean difference (SMD) and 95% confidence intervals (CIs). We rationalised data presentation between studies with consistent direction of effect. Where scales differed for outcomes, we calculated SMDs instead, with corresponding 95% CIs. Using previously described methods we translated SMD to a typical scale (e.g. 0–10 for pain) by multiplying the SMD by a typical among-person SD (e.g. the SD of the control group at baseline from the most representative trial).⁷⁷ We analysed dichotomous data as risk ratios (RRs) or Peto's odds ratio (OR) when the outcome was a rare event (approximately < 10%), and use 95% CIs.

Where a single trial reported multiple time points for the same outcome, then we extracted the data that related to the latest time point within each follow-up category (e.g. if one trial reports outcomes at 6 months and 1 year, we extracted the 1-year results only).

Dealing with missing data

Where possible we obtained missing data from authors or research sponsors. In cases where we could not collect further data, the effect of including the work was investigated through sensitivity analysis (SA). The same process was used to assess the effect of studies with loss to follow-up of over 20%. Studies with > 50% attrition were excluded from the meta-analysis.

Where possible, we computed missing SDs from other statistics, such as standard errors (SEs), CIs or p -values.⁷⁸ If SDs could not be calculated, they were imputed (e.g. from other studies in the meta-analysis).

Assessment of reporting biases

We planned to create and examine funnel plots to explore possible small-study biases for outcomes where 10 or more studies are included in a meta-analysis. In interpreting funnel plots, we examined the different possible reasons for funnel plot asymmetry and relate this to the results of the review.⁷⁹ We evaluated whether selective reporting of outcomes was presented by comparing included study reports with their registered study protocol when available through the ClinicalTrials.gov or World Health Organization International Clinical Trials Registry platforms.

Subgroup analysis and investigation of heterogeneity

Where sufficient comparative data were available, we planned to carry out the following subgroup analyses for outcomes related to patient-reported pain and function:

1. Age under 40, 40–49, 50–59, 60 years and over.
2. Sex.
3. Grade of structural disease determined by KL grade (KL1–3 and KLO, i.e. morphological features but no radiological change).
4. Surgical approach (for osteochondroplasty of the hip only).

Sensitivity analysis

Where sufficient data were available, we planned to carry out SAs examining the effects of quality criteria:

1. restricting to studies at low or moderate overall risk of bias
2. different analytical approaches to handling missing outcome data (e.g. ITT vs. available case analysis)
3. including studies with an incomplete description of EKO or EHO.

Summary of findings tables

We produced 'summary of findings' tables (showing the results separately for hip and knee procedures) based on each of the four main groups of surgical interventions (non-regenerative procedures, biological reconstruction, load-modifying procedures and joint replacement) compared to current non-operative management following NICE guidance.

We used the GRADE approach using GRADEpro (Evidence Prime Inc, Hamilton, Ontario, Canada) software (i.e. risk of bias, inconsistency of effect, imprecision, indirectness and publication bias) to assess the quality of a body of evidence. If there was no evidence from comparative studies for a particular comparison, we showed blank rows in the summary of findings table. We did not conduct a GRADE assessment on studies that did not compare surgery with non-operative care.

Overviews of reviews of non-comparative studies to inform research

If there is inadequate evidence from comparative studies, we conducted additional analysis of non-comparative studies as additional analysis. Such studies cannot provide effect sizes for the clinical and cost-effectiveness analyses; however, they can be used to identify interventions used in practice that should be evaluated in future trials.

Data to be taken forward to the health economic analysis

We planned that outcome data from studies where a type of surgery was compared to non-operative treatment could be taken forward to the health economic analysis. We included all comparative studies including RCT studies and NRS of surgery compared against non-operative treatment. Data from RCTs or NRS of one type of surgery compared to another type of surgery, used in the treatment of early hip or knee OA, could be not used to provide data for the health economic analysis. However, the RCTs are reported to offer some evidence of the range of surgical treatments used and their outcome.

Time-line and adjustments to protocol

Following the COVID lock-down, this study received a time extension. The write-up period has taken a long time due to the complexity of the work. The searches for the main systematic review were completed on January 2021. After discussion with the HTA it was agreed that additional search for RCTs would be repeated so any new studies could be added. The search was completed in January 2025 using ClinicalTrials.gov and ISRCTN.com. No new RCT studies of the surgical treatment of early hip or knee arthritis compared to non-operative treatment were identified.

Chapter 3 Non-regenerative procedures used to treat early osteoarthritis in the knee

Results of the search

The search and successive updates up to January 2021 generated 10,683 records. After deduplication and first assessment based on title and abstract screening, we obtained 134 in full text, comprising 26 systematic reviews of interventions and 113 potential RCTs or comparative NRS ([Figure 1](#)). The search strategy is recorded in [Appendix 1](#).

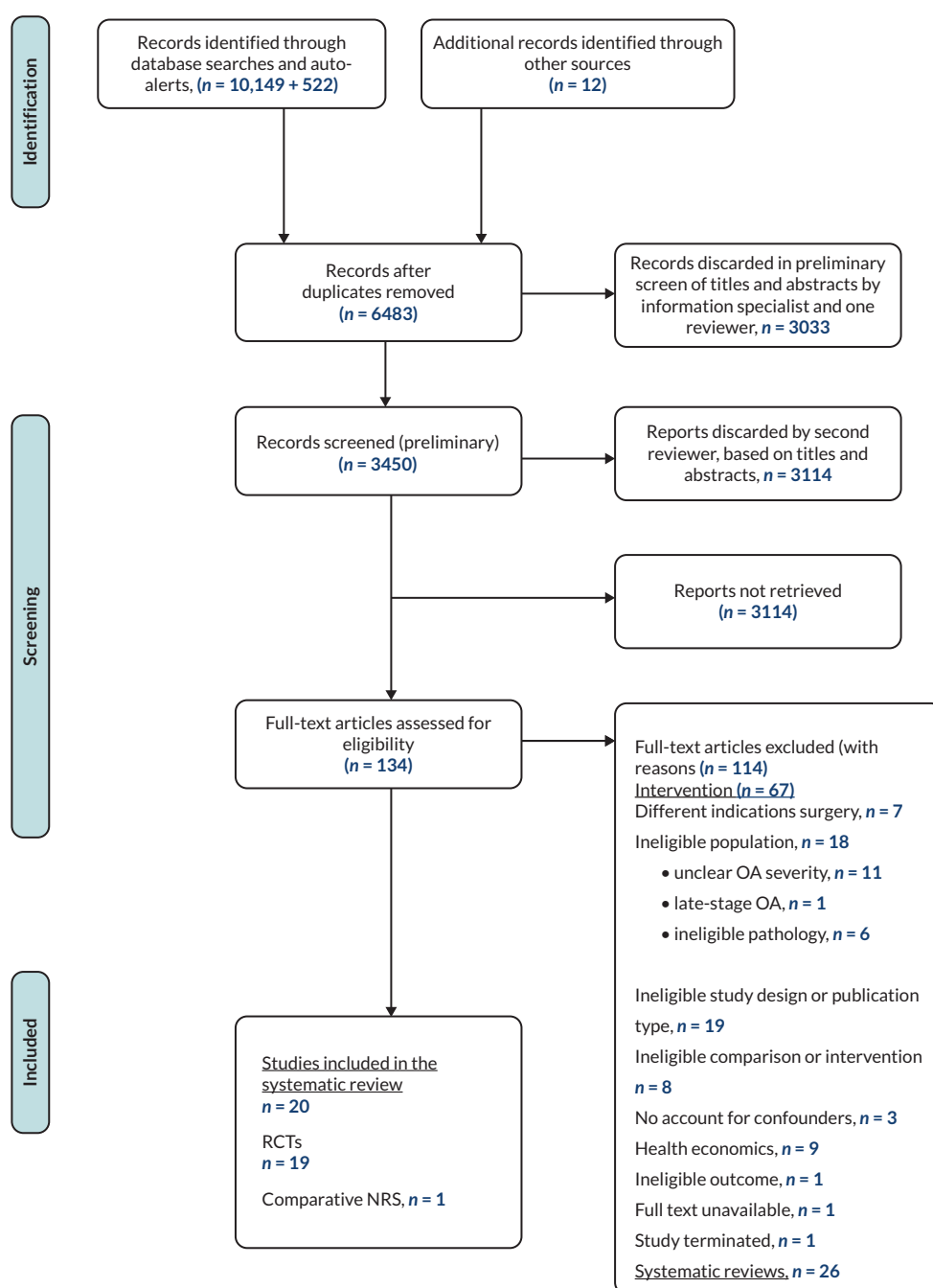


FIGURE 1 PRISMA diagram: non-regenerative surgery for treatment of early knee OA.

Finally, 19 RCTs and 1 comparative NRS were eligible and included in the analysis.^{21,22,80-97} We excluded 93 studies. The main reasons for exclusion were an ineligible study design or publication type, an ineligible population, in 11 of which the OA severity of participants was unclear, and a number systematic reviews were excluded on the basis that they could not be used as sole evidence without modification. Other reasons for exclusion included ineligible interventions or comparators, NRS that took no account for confounding factors, ineligible outcome measures and studies which pertained to health economics outcomes.

Through the search we identified a number of important NICE guidelines that have been published relating to surgical treatment of early OA.⁹⁸

No ongoing studies were identified at the time of this review.

Included studies characteristics

The main features of the characteristics of included studies are summarised in [Tables 2](#) and [3](#).

Study details

A total of 22 studies were included in the review. There were 19 RCTs and 3 non-randomised controlled studies.

Characteristics of included studies

The studies included can be divided into two groups. The first group involves comparative studies where the surgical intervention was used directly to treat people with early OA (as defined in our protocol) (see [Table 2](#)). We included seven RCTs and one NRS in this group.^{21,81,82,85,86,88,92,94} Data from this group of studies were eligible to be taken forward to the health economic analysis.

The second group included studies where the surgical intervention used to treat specific related knee conditions for people in whom some or all participants have early OA (indirect treatment of early OA) (see [Table 3](#)). This included 12 RCTs and 2 NRSs in this group, one of which had been reported twice in the literature.^{22,80,83,84,87-91,93,95-97,99-101}

The eligible RCTs involved 2529 participants were included in this review.^{21,22,82-93,95-97} For these studies, the median sample size was 114 [interquartile range (IQR) 78–165]; 11 RCTs had more than 100 participants with the largest containing 351.^{21,22,80,83,87-89,92,95-97} Pre-published protocols were available for seven of the trials.^{22,80,83,87,89,91,95} In addition, three NRS were also included with sample sizes ranging from 670 to 4234.^{94,99-101}

Of the 19 RCTs, 2 were conducted in the UK 5 in the USA, 2 in Sweden and single RCTs were completed in Canada, Denmark, Finland, Germany, India, Korea, Norway, Romania, Spain and The Netherlands.^{21,22,80-87,89-93,95-97} The NRS were conducted in China and the USA.^{94,99-101}

Seven were multicentre trials, with the largest being conducted over 9 sites, 2 were conducted across 2 sites and 12 were single-centre studies.^{21,22,80-87,89-93,95-97}

Two studies had industry funding.^{80,90} Six were supported by public funding sources.^{21,22,85,87-89} A further three received a mixture of industry and non-industry funding.^{81,91,95} The remaining eight RCTs did not report any funding source.^{82-84,86,92,93,96,97}

Population, intervention, control/comparison, outcome and participant characteristics

These are recorded in [Tables 2](#) and [3](#).

Surgical interventions assessed

The surgical interventions performed are presented as three main three types of procedure; type 1 studies: comparisons of non-regenerative surgical procedures versus non-operative management; type 2: comparisons of one non-regenerative surgical procedure versus another non-regenerative surgical procedure; and type 3:

TABLE 2 Summary of included non-regenerative surgery studies characteristics for direct treatment of early knee OA

Study, design and country	Study size and duration	Population characteristics: mean age (SD), sex; male/female (%) BMI (kg/m ²), symptom duration (months)	Indication and stage of OA	Intervention	Comparator	Outcomes
Type 1: comparisons of non-regenerative surgical procedures vs. non-operative management for directly treating early knee OA (RCTs and NRS)						
Chang 1993 ⁸¹ RCT (2 centres) USA Direct OA treatment	32 (18 and 14) 1 year	<ul style="list-style-type: none"> Age: 61 (11) and 65 (13) years Sex: 9 (28%) and 10 (29%) BMI: 29.1 (5.1) and 30.0 (4.6) Symptoms: mean 52 months 	<ul style="list-style-type: none"> Early OA KL1: 22% and 14%; KL2: 28% and 36%; KL3: 50% and 50% 	Arthroscopic debridement (included debridement, lavage and excision of loose cartilage) + physiotherapy (strengthening and flexibility exercises and gait training)	Closed needle joint lavage + physiotherapy	Pain, function, pain and function
Forster 2003 ⁸² RCT (single-centre) UK Direct OA treatment	38 (19 and 19) 1 year	<ul style="list-style-type: none"> Age: 63 and 60 (SD NR) years Sex: NR BMI: NR Symptoms: NR 	<ul style="list-style-type: none"> Unclear if early OA KL grade NR; however, authors refer to 'some remaining joint space'. 	Arthroscopic lavage [Debridement of the articular surface or menisci undertaken as considered clinically necessary (but majority only had a washout)].	HA injection	Combined pain and function, pain, function, reoperation, TKA
Su 2018 ⁹⁴ NRS Direct OA treatment	382 5 years	<ul style="list-style-type: none"> Age: 56 (15) and 57 (14) Sex 36 (20%) and 28 (20%) BMI: 28 (5) and 27 (5) 	<ul style="list-style-type: none"> KL 2 22% KL 3 57% KL 4 21% 	Arthroscopic debridement and debridement	Physiotherapy and exercise IA injection	Pain, function, TKA
Kirkley 2008 ⁸⁸ RCT (single-centre) Canada Direct OA treatment	188 (94 and 94) 2 years	<ul style="list-style-type: none"> Age: 58.6 (10.2) and 60.6 (9.9) years Sex: 38 (41%) 28 (33%) BMI: 31.6 (6.7) and 30.2 (6.3) Symptoms: 47.1 (69.4) and 40.1 (72.6) 	<ul style="list-style-type: none"> Mixed stage OA with subgroup data available KL2: 44% KL3: 51% KL4:5% 	Arthroscopic debridement (including some meniscal excision) + physiotherapy/HEP	Physiotherapy/ Home exercise program	Combined pain and function, function, pain, QoL
Saeed 2015 ⁹² RCT (single-centre) India Direct OA treatment	120 (60 and 60) 6 months	<ul style="list-style-type: none"> Age: 18% 40–45 years; 21% 46–50 years; 27% 56–60 years; 34% over 60 years (NR for each arm of the trial) Sex: 12 (28%) entire study population BMI: NR Symptoms: NR 	<ul style="list-style-type: none"> Early OA KL ≤ 3 both groups 	Arthroscopic debridement	HA injection	Pain
Type 2: comparisons of one non-regenerative surgical procedure vs. another non-regenerative surgical procedure for directly treating early knee OA (RCTS only)						
Hubbard 1996 ⁸⁵ RCT (single-centre) UK Direct OA treatment	76 (40 and 36) 5 years	<ul style="list-style-type: none"> Age: 51 and 45 'success' and 'fail' groups for intervention 46 and 59 for comparator Sex: 28 (70%) and 20 (56%) BMI: NR Symptoms: NR 	<ul style="list-style-type: none"> Unclear if early OA KL grade NR Single medial femoral condyle degenerative lesion grade 3 or 4 	Arthroscopic debridement All loose articular cartilage was resected; no attempt was made to abrade the bone or to drill the condyles	Arthroscopic lavage	Pain

Study, design and country	Study size and duration	Population characteristics: mean age (SD), sex; male/female (%) BMI (kg/m ²), symptom duration (months)	Indication and stage of OA	Intervention	Comparator	Outcomes
Kalunian 2000 ⁸⁶ RCT (4 centres) USA <i>Direct OA treatment</i>	90 (41 and 49) 1 year	<ul style="list-style-type: none"> Age: 61 and 58 years (41–88 range for both groups, SD NR) Sex: 19 (46%) and 23 (47%) BMI: NR Symptoms: NR 	<ul style="list-style-type: none"> Early OA KL ≤ 3 in both groups 	Arthroscopic lavage (3 l)	Arthroscopic lavage (250 ml)	Combined pain and function, pain, function
Moseley 2002 ²¹ RCT (single-centre) USA <i>Direct OA treatment</i>	180 (59 AD, 61 AL, 60 sham) 2 years	<ul style="list-style-type: none"> Age: AD: 54 (12.2), AL: 51 (10.5) and Sham: 52 (11.1) years Sex: AD: 96.6%, AL: 88.5% and Sham: 93.3% BMI: NR Symptoms: NR 	<ul style="list-style-type: none"> ~80% with KL ≤ 3 Mild 28% (1–3) Moderate 46% (4–6) Severe 26% (7–8) (sum of KL grade 0–4 for each compartment) 	Arthroscopic debridement	Arthroscopic lavage Sham surgery	Pain, function, QoL
Type 3: comparisons of non-regenerative surgical procedures vs. other joint-preserving or joint replacement surgical options for directly treating early knee OA						
No studies						
HA, hyaluronic acid.						

TABLE 3 Summary of included non-regenerative surgery studies characteristics for indirect treatment of early knee OA

Study, design and country	Study size and duration	Population characteristics: mean age (SD), sex; male/female (%) BMI (kg/m ²), symptom duration (months)	Indication and stage of OA	Intervention	Comparator	Outcomes
<i>Type 1: comparisons of non-regenerative surgical procedures vs. non-operative management for indirectly treating early knee OA (RCTs and NRS)</i>						
Gauffin 2014 ⁸³ RCT (single-centre) UK <i>Indirect OA treatment</i>	150 (75 and 75) 5 years	<ul style="list-style-type: none"> Age: 54 (5) and 54 (6) years Sex: 53 (71%) and 56 (75%) BMI: NR Symptoms: 7 (8) and 7 (7) median/IQR 	<ul style="list-style-type: none"> Meniscal pathology KL0: < 50%, remaining study population KL1-2 	APM + physiotherapy	Home exercise programme	Pain, function, HRQoL, radiological outcomes, reoperations, AEs
Herrlin 2007 ⁸⁴ RCT (single-centre) Sweden <i>Indirect OA treatment</i>	99 (47 and 49) 5 years	<ul style="list-style-type: none"> Age: 54 (5) and 56 (6) median/IQR Sex: 28 (60%) and 27 (63%) BMI: 26 (3.9) and 26 (4.9) Symptoms: NR 	<ul style="list-style-type: none"> Meniscal tear Ahlback ≤ 1 	APM	Physiotherapy	Combined pain and function, pain, function, HRQoL, radiological outcomes, reoperation
Katz 2013 ⁸⁷ (METEOR) RCT (7 centres) USA <i>Indirect OA treatment</i>	351 (174 and 177) 5 years	<ul style="list-style-type: none"> Age: 58.6 (7.9) and 57.2 (6.7) years Sex: 75 (43%) and 75 (42%) BMI: 30.2 (6.2) and 30.2 (6.1) Symptoms: NR 	<ul style="list-style-type: none"> Meniscal tear KL ≤ 3 both groups 	APM + physiotherapy + Home exercise programme	Home exercise programme + physiotherapy	Pain, function, HRQoL, TKA, AEs, SAEs
Kise 2016 ⁸⁹ RCT (2 centres) Norway <i>Indirect OA treatment</i>	140 (70 and 70) 5 years	<ul style="list-style-type: none"> Age: 49 (6) and 50 (6) years Sex: 43 (61%) and 43 (61%) BMI: 26 (3.7) and 26.4 (4.3) Symptoms: 12 (15.7) and 17.3 (21.5) 	<ul style="list-style-type: none"> Meniscal tear 50% or more KLO; rest KL1-2 	APM	Physiotherapy	Combined pain and function, pain, function, HRQoL, reoperation, radiological outcomes, AEs
Merchan 1993 ⁹⁰ RCT (single-centre) Spain <i>Indirect OA treatment</i>	80 (40 and 40) 3 years	<ul style="list-style-type: none"> Age: 57 (50-63 range) and 57 (50-65 range) Sex: 7 (20%) and 13 (34%) BMI: NR Symptoms: NR 	<ul style="list-style-type: none"> Meniscal pathology Limited degenerative process (minimal JSN and formation of small osteophytes) 	APM + home exercise programme + physiotherapy	NSAIDs + conservative management including physiotherapy	Combined pain and function, AEs (mortality)
Van de Graaf 2016 ⁹⁵ (ESCAPE) RCT (9 centres) The Netherlands <i>Indirect OA treatment</i>	321 (159 and 162) 2 years	<ul style="list-style-type: none"> Age: 58 (7) and 57 (7) years Sex: 78 (49%) and 80 (50%) BMI: 26.7 (3.8) and 27.2 (4) Symptoms: NR 	<ul style="list-style-type: none"> Meniscal tear < 50% KLO; rest KL1-2KL 	APM + home exercise programme	Physiotherapy	Pain, function, progression of OA, HRQoL, reoperations, AEs, TKA
Yim 2013 ⁹⁷ RCT (single-centre) Korea <i>Indirect OA treatment</i>	108 (54 and 54) 2 years	<ul style="list-style-type: none"> Age: 55 (10) and 58 (11) years Sex: 9 (18%) and 12 (23%) BMI: 25 (2.5) and 26.4 (1.9) Symptoms: NR 	<ul style="list-style-type: none"> Meniscal tear KL ≤ 3 + pain both groups 	APM	Physiotherapy	Pain and function, radiological outcomes
Vermesan 2013 ⁹⁶ RCT (single-centre) Romania <i>Indirect OA treatment</i>	114 60 and 60 (knees) 1 year	<ul style="list-style-type: none"> Age: 59 (8) and 58 (8) years Sex: 11 (18%) and 14 (77%) BMI: 32.7 (6.4) and 31.9 (6.2) Symptoms: 3 (1.5) and 3 (1.7) 	<ul style="list-style-type: none"> Medial compartment OA Authors state 'early stage' OA 	APM/debridement	IA steroid injection	Combined pain and function, TKA, reoperation/repeat treatment

Study, design and country	Study size and duration	Population characteristics: mean age (SD), sex; male/female (%) BMI (kg/m ²), symptom duration (months)	Indication and stage of OA	Intervention	Comparator	Outcomes
Type 2: comparisons of one non-regenerative surgical procedure vs. another non-regenerative surgical procedure in the knee (including injection) (RCTs only)						
Bisson 2017 ⁸⁰ (CHAMP) RCT (1 centre) USA <i>Indirect OA treatment</i>	190 (98 and 92) 1 year	<ul style="list-style-type: none"> Age: 55 (7.3) and 54 (7.9) years Sex: 121 men (64%) and 59 (64.1%) BMI: 29.1 (5.1) and 30.0 (4.6) Symptoms: duration NR 	<ul style="list-style-type: none"> Meniscal tear and chondral lesions KLO or 1 	APM plus chondroplasty with motorised shaver	APM without chondral debridement	Pain, physical function, pain and function, HRQoL, AEs
Roos 2018 ⁹¹ RCT (3 centres) Denmark <i>Indirect OA treatment</i>	44 (22 and 22) 2 years	<ul style="list-style-type: none"> Age: 47 (6) and 46 (6) years Sex: 13 (59%) and 10 (45%) BMI: 27.6 (3.6) and 26 (3.9) Symptoms: 5 (2–6) and 3.5 (2–6) median/IQR 	<ul style="list-style-type: none"> Meniscal tear < 50% KLO; rest KL1–2 	APM + home exercise programme	Sham surgery + home exercise programme	Combined pain and function, pain, function, HRQoL, reoperation, AEs
Sihvonen 2013 ²² (FIDELITY) RCT (5 centres) Finland <i>Indirect OA treatment</i>	146 (70 and 76) 2 years	<ul style="list-style-type: none"> Age: 52 (7) and 52 (7) years Sex: 42 (60%) and 47 (62%) BMI: 26.9 (4) and 27.9 (4) Symptoms: 10 (6.6–14) and 9.7 (6.4–16) median/IQR 	<ul style="list-style-type: none"> Meniscal tear < 50% KLO; rest KL1–2 	APM	Sham surgery	Combined pain and function, pain, HRQoL, AEs
Spahn 2008 ⁹³ RCT (single-centre) <i>Indirect OA treatment</i>	60 (30 and 30) 10 years	<ul style="list-style-type: none"> Age: 42.9 (10.4) and 43.7 (9.0) years (intervention and comparator) Sex: 13 male (43.3%) and 15 male (50%) BMI: 26.6 (5.1) and 27.8 (4.3) Symptoms: 8.4 and 7.4 	<ul style="list-style-type: none"> Medial meniscus tear and an idiopathic Grade III cartilage defect(s) of the medial femoral condyle ICRS grade III defect of the medial femoral condyle 	APM followed by bipolar radiofrequency chondroplasty (RFC)	APM followed by MSD	Combined pain and function, pain, function, HRQoL, progression, reoperation, AEs
Type 3: comparisons of non-regenerative surgical procedures vs. other joint-preserving or joint replacement surgical options in the knee.						
No studies						
APM, arthroscopic partial meniscectomy; MSD, mechanical shaver debridement; SAE, serious adverse event.						

comparisons of non-regenerative surgical procedures versus other joint-preserving or joint replacement surgical options. For type 1 procedures we have grouped the type of surgical intervention into two main categories, 'direct' treatment of early OA and 'in-direct' where surgical intervention used to treat specific knee conditions in patients, where some or all participants have early OA.

The following surgical interventions were reported in the included studies:

1. Direct treatment of early OA – seven RCTs and one NRS in this group.^{21,81,82,85,86,88,92,94}
 - Arthroscopic debridement: five RCTs and one NRS.^{21,81,85,88,92,94}
 - Arthroscopic lavage: three RCTs.^{21,82,86}
2. Indirect treatment of early OA – 12 RCTs and 2 NRSs in this group, one of which had been reported twice in the literature.^{22,80,83,84,87,89–91,93,95–97,99–101}
 - Arthroscopic partial meniscectomy (APM): nine RCTs and two NRS.^{22,80,83,84,87,89,91,95,97,99–101}
 - APM plus chondral debridement: two RCTs.^{80,93}
 - APM plus bipolar radiofrequency chondroplasty (RFC): one RCT.⁹³

Three RCTs compared two of these surgical interventions.^{80,85,93}

Comparator interventions assessed

The following non-operative comparators were reported in the included studies:

- Physiotherapy: two RCTs.^{89,95}
- Home exercise programme: two RCTs.^{83,84}
- Combination of physiotherapy and home exercise programme: three RCTs.^{87,88,97}
- IA hyaluronic acid (HA) injection (as a class): two RCTs.^{82,92}
- IA steroid injection: one RCT.⁹⁶
- Tidal lavage: one RCT.⁸¹
- No arthroscopy: one NRS.⁹⁴
- No APM: two NRS.^{99–101}
- Combination treatment (oral medications, physiotherapy, HA or steroid injection): one RCT and one NRS.^{90,94}

The following control surgical comparators were reported:

- Sham surgery (as a surgical comparator): three RCTs.^{21,22,91,92}
- Arthroscopic minimal irrigation (250 ml): one RCT.⁸⁶

No studies explicitly reported NICE 'core' guidance as a comparator (combination of information, exercise/physical therapy modalities and weight loss) or the individual components.

Outcomes

All studies reported at least one major outcome described in our protocol to be eligible. Most outcomes related to pain or physical function, combined pain and function, and HRQoL used at least one validated instrument. Most studies only reported results at short-term follow-up (≤ 2 years); only four RCTs and one NRS reported any results at medium-term follow-up (> 2 years to < 5 years) and six RCTs and three NRS reported at long-term follow-up (5 years or more).^{83–85,87,89,90,93,94,99–101}

The following outcomes were reported:

- Pain: 16 RCTs and 1 NRS.^{21,22,80–84,86–89,91–95,97} (pain as a dichotomous outcome: 3 RCTs)^{83,85,97}
- Physical function: 13 RCTs.^{21,80–84,86–89,91,93,95}
- Pain and function: 11 RCTs and 1 NRS.^{22,81,82,84,86,88,89,91,93,94,96,97} (pain and function as a dichotomous outcome: 2 RCTs.)^{89,90}

- PREMs: not reported
- HRQoL: 11 RCTs.^{21,80,83,84,87–89,91,93}
 - Whole person QoL scales: EQ-5D: two RCTs;^{83,91} Short Form questionnaire-36 items (SF-36) physical function: five RCTs.^{21,80,87,88,95}
 - Knee specific scales, KOOS QoL: four RCTs,^{84,89,91,93} WOMET: one RCT,²²
- Surgical complications and AEs: 11 RCTs and 1 NRS.^{22,80,83,85,87–91,93–95}
- Joint replacement (2 years' follow-up or more): five RCTs and three NRS.^{22,87,88,93–95,99–101}
- Reoperation rate/revision of initial surgical procedure, including any new procedure, and time to revision or other reoperation: 11 RCTs and 1 NRS.^{22,82–84,87–89,91,93–96}
- Radiological outcomes (progression of OA – 2 years' follow-up or more): five RCTs and one NRS.^{83,84,89,93,97,100,101}

Results

We present the results in three main sections reflecting the different studies; type 1: comparisons of non-regenerative surgical procedures versus non-operative management, type 2: comparisons of one non-regenerative surgical procedure versus another non-regenerative procedure and type 3: comparisons of a non-regenerative procedure versus another surgical procedure. Within each category, where available, we report each outcome at short-, medium- and long-term follow-up for each of the individual comparisons, giving results separately for RCTs and NRS. The findings are summarised in [Table 4](#). Risk-of-bias assessment for included studies is presented in [Appendix 2](#).

The results are split into two sections depending whether the surgery was direct treatment of early OA, or indirect treatment of early OA (used to treat specific knee conditions for people in whom some or all participants have early OA).

Within each category, where available, we report each outcome at short-, medium- and long-term follow-up for each of the individual comparisons (e.g. arthroscopic debridement vs. physiotherapy), giving results separately for RCTs and NRS. Where the comparator was physiotherapy or a home exercise programme or a combination of these, we defined a common comparator, 'physiotherapy'. Additionally, we treated comparisons of 'surgical intervention plus physiotherapy versus physiotherapy' as 'surgical intervention versus physiotherapy' as some form of physiotherapy is part of standard care for all participants undergoing surgery for OA.

Only results for type 1 studies where a surgical treatment was used directly to treat early knee OA could be taken forward to the health economic analysis.

Direct treatment of early osteoarthritis

Type 1 comparisons of non-regenerative surgical procedures versus non-operative management to directly treat early knee osteoarthritis.

Four RCTs in 380 randomised participants and 1 comparative NRS in 382 participants reported the effect of non-regenerative surgical procedures in comparison with non-operative management.^{81,82,88,92,94} Not all studies reported all the outcomes (see [Outcomes](#)). No studies reported progression of OA or any PREMs outcomes.

The surgery versus non-operative management comparisons in this section were too heterogeneous by design to pool within a meta-analysis, so we report the results separately for each comparison. To compare across different comparisons or across different time points, where there are different scales, we report the SMD with its 95% CI for single studies and give the mean difference in brackets. Results are reported for RCTs unless otherwise stated. Where studies reported PROMs outcomes on a positive scale (high values indicating good outcomes), we converted the results to a negative scale (indicated with an asterisk). Therefore, a negative SMD value routinely means that the effect favours surgery, with the exception of HRQoL (in which case a positive value favours surgery).

TABLE 4 Summary of findings table for non-regenerative surgery studies for treatment of early knee OA

Non-regenerative surgical procedure compared with non-operative management for early OA of the knee						
Patients or population: people with early OA of the knee Intervention: non-regenerative surgical procedure				Settings: primary and secondary care Comparison: non-operative management		
Continuous outcomes	Score with non-operative management	Mean difference (back-transformed)	SMD (95% CI) findings	Sample size (no. studies)	Evidence certainty (GRADE)	Comments
<i>Direct treatment of early knee OA. RCT evidence: arthroscopic debridement (which included some limited meniscal excision) vs. physiotherapy. NRS evidence: arthroscopic debridement (no APM, some MF) vs. mixed conservative treatment (physiotherapy, medication, HA acid injection, steroid injection)</i>						
Pain at 2 years	WOMAC pain score (0–500) = 185 (SD 132) (baseline: 214)	MD: –17.0 (–57.3 to 23.3)	SMD: –0.13 (–0.43 to 0.18)	168 (1 RCT)	⊕⊕⊕⊕ Very low ¹	RCT. Pain analysed as a negative outcome Negative SMD and MD favours surgery Improvement in WOMAC pain over baseline for non-operative group = 29 (scale 0–500)
<i>Pain at long-term duration – not measured</i>						
Physical function (impairment) at 2 years (short-term)	WOMAC functional subscale (0–1700) 623 (439) (baseline 726)	MD: –11.00 (–145.2 to 123.2)	SMD –0.02 (–0.33 to 0.28)	168 (1 RCT)	⊕⊕⊕⊕ Very low ¹	Function analysed as a negative outcome. Negative SMD and MD favours surgery Improvement in WOMAC physical function over baseline for non-operative management group = 103 (scale 0–1700)
<i>Physical function at long-term duration – not measured</i>						
Pain and function at 2 years (short-term)	WOMAC total score (0–2400) = 897 (SD 583) (baseline 1043)	MD: –23.0 (–206 to 160) Large baseline difference 144 points masks effect	SMD: –0.04 (–0.34 to 0.26)	168 (1 RCT)	⊕⊕⊕⊕ Very low ¹ /critically low	RCT. Large baseline difference of 144 WOMAC points (surgery group had poorer function) Improvement in WOMAC total score over baseline for non-operative management group = 146 (scale 0–2400)
	WOMAC total score (0–96) KL2 = 33 (SD 13) KL3 = 34 (SD 13)	MD (favouring surgery) KL2: –5.3 (–10.2 to –0.4) KL3: –7.1 (–10.4 to –3.9)	SMD KL2: –0.50 (–0.96 to –0.04) KL3: –0.60 (–0.88 to –0.32)	KL2 = 75 KL3 = 211 (1 NRS)	⊕⊕⊕⊕ Very low ²	NRS. Analysis was adjusted for baseline and stratified by KL grade. Participants refusing surgery assigned to conservative treatment – likely selection bias.
Pain and function at 5 years (long-term)	WOMAC total (0–96) KL2 = 30 (SD 11) KL3 = 37 (SD 15)	MD (favouring surgery) KL2: + 0.8 (–5.5 to 7.1) KL3: –3.4 (–7.6 to 0.8)	SMD KL2: + 0.07 (–0.47 to 0.60) KL3: –0.31 (–0.69 to 0.07)	KL2 = 55 KL3 = 109 (1 NRS)	⊕⊕⊕⊕ Very low ²	NRS; details as above. Number of participants estimated. Negative SMD favours surgery
QoL, 2 years (short-term)	SF-36 physical component score (0–100) = 37 (SD 11) (baseline 34)	MD: –0.20 (–3.53 to 3.13)	–0.02 (–0.32 to 0.28)	168 (1 RCT)	⊕⊕⊕⊕ Very low ¹	RCT. QoL analysed as a positive outcome. Positive MD favours surgery. Improvement in SF-36 physical component over baseline for non-operative management group = 3 (scale 0–100)

TABLE 4 Summary of findings table for non-regenerative surgery studies for treatment of early knee OA (*continued*)

Non-regenerative surgical procedure compared with non-operative management for early OA of the knee						
Patients or population: people with early OA of the knee Intervention: non-regenerative surgical procedure			Settings: primary and secondary care Comparison: non-operative management			
Continuous outcomes	Score with non-operative management	Mean difference (back-transformed)	SMD (95% CI) findings	Sample size (no. studies)	Evidence certainty (GRADE)	Comments
Binary outcomes	Absolute risk non-operative	Absolute risk difference (95% CI)	Relative effect (95% CI)	Number of participants (studies)	Evidence certainty (GRADE)	Comments
TKR 2, and 5 years' follow-up	At 2 years: 14 people per 1000 had a TKA	2 years: 8 fewer people had a TKA per 1000 (from 13 fewer to 52 more)	2-year follow-up: RR = 0.43 (0.04–4.68)	301 (1 NRS) 2 years: 3 TKR	⊕⊕⊕⊕ Very low ³	NRS. Not adjusted analysis. Results combined for KL grades 2 and 3 Participants refusing surgery assigned to conservative treatment – likely selection bias. Negative outcome, RR < 1 favours surgery
	At 5 years: 101 people per 1000 had a TKA	5 years: 33 fewer people had a TKA per 1000 (from 69 fewer to 44 more)	5-year follow-up: RR = 0.67 (0.32–1.44)	5 years: 25 TKR		
Reoperation rate: outcome not reported for comparison of arthroscopic debridement vs. physiotherapy						
AEs: one NRS in 382 participants reported no AEs in either arm up to 5 years						
Progression: no studies reported this outcome						
MD, mean difference.						

We will report the findings of the RCT evidence only as NRS are not eligible to take forward to the health economic analysis.

Pain at short-term follow-up

- Arthroscopic debridement (including some meniscal excision) versus physiotherapy:⁸⁸
 - Short-term (2 years): SMD = -0.13 (-0.43 to 0.18). Very low certainty evidence; 1 RCT in 168 participants [MD = -17.0 (-57.3 to 23.3), WOMAC pain 0–500 scale]
- Arthroscopic lavage versus IA-HA injection:⁸²
 - Short-term (12 months): SMD = 0.00 (-0.79 to 0.79). Very low certainty evidence; 1 RCT in 25 participants [MD = 0.00 (-2.47 to 2.47), visual analogue scale (VAS) 0–10 scale; control group means: baseline 7.6, postoperative 5.7].
- Arthroscopic debridement versus IA-HA injection:⁹²
 - Short-term (6 months): SMD = 0.92 (0.54 to 1.29), favouring HA. Very low certainty evidence; 1 RCT in 120 participants [MD = 6.80 (4.16 to 9.44), Knee Society Score (KSS) pain scale 0–50; control group means: baseline × 27 (= 50 minus reported baseline), postoperative × 12].
- Arthroscopic debridement versus tidal irrigation:⁸¹
 - Short-term (12 months): SMD = 0.15 (-0.55 to 0.85). Very low certainty evidence; 1 RCT in 32 participants [MD = 0.30 (-1.10 to 1.70), AIMS pain scale 0–10; control group means: baseline 6.1, postoperative 5].

Physical function at short-term follow-up

- Arthroscopic debridement (including some meniscal excision) treatment versus physiotherapy:⁸⁸
 - Short-term (2 years): SMD = -0.02 (-0.33 to 0.28). Very low certainty evidence; 1 RCT in 168 participants (large baseline difference of 104 WOMAC points; surgery participants had worse physical function at baseline) [MD = -11.00 (-145.2 to 123.2), WOMAC function scale 0–1700].
- Arthroscopic lavage versus IA-HA injection:⁸²
 - Short-term (12 months): SMD = 0.72 (-0.10 to 1.54), that is towards less effectiveness for surgery. Very low certainty evidence; 1 RCT in 25 participants (large baseline difference of 22.4 KSS points; surgery participants had worse physical function at baseline) [MD = 23.9 (-1.45 to 49.3), KSS function 0–100].
- Arthroscopic debridement versus tidal irrigation:⁸¹
 - Short-term (12 months): SMD = -0.27 (-0.97 to 0.44). Very low certainty evidence; 1 RCT in 32 participants (very large baseline difference of 0.6 AIMS points; surgery participants had worse physical function at baseline; results adjusted for the baseline) [MD = -0.30 (-1.07 to 0.47), AIMS function scale 0–10].

Pain and function at short-, medium- and long-term follow-up

- Arthroscopic debridement (including some meniscal excision) versus physiotherapy:⁸⁸
 - Short-term (2 years): SMD = -0.04 (-0.34 to 0.26). Very low certainty evidence; 1 RCT in 168 participants (large baseline difference of 144 WOMAC points; surgery participants had worse total WOMAC scores at baseline) [MD = -23.0 (-206 to 160), WOMAC scale 0–2400].
- Arthroscopic lavage versus IA-HA injection:⁸²
 - Short-term (12 months): SMD = 0.42 (-0.38 to 1.22). Very low certainty evidence; 1 RCT in 25 participants [MD = 2.90 (-2.30 to 8.10), Lequesne index 0–24].
- Arthroscopy treatment (some MF included but no APM) versus conservative treatment (physiotherapy, medication, HA injection and steroid injection):⁹⁴
 - Short-term (2 years): KL grade 2, SMD = -0.50 (-0.96 to -0.04), favouring surgery. One RCT in 75 participants. KL grade 3, SMD = -0.60 (-0.88 to -0.32); 211 participants. Very low certainty evidence [MD = -5.3 (-10.2 to -0.4) for KL2 and -7.1 (-10.4 to -3.9) for KL3, WOMAC scale 0–96].

- Medium-term (4 years): KL grade 2: SMD = 0.46 (−0.05 to 0.97); 62 participants. KL grade 3: SMD = −0.20 (−0.48 to 0.08); 1 RCT in 198 participants. Very low certainty evidence [MD = 4.50 (−0.23 to 9.23) for KL2 and −2.80 (−6.75 to 1.15) for KL3, WOMAC scale 0–96].
- Long-term (5 years): KL grade 2: SMD = 0.07 (−0.47 to 0.60); 55 participants. KL grade 3: SMD = −0.31 (−0.69 to 0.07); 1 RCT in 109 participants (substantial loss to follow-up). Very low certainty evidence [MD = 0.8 (−5.5 to 7.1) for KL2 and −3.4 (−7.6 to 0.8) for KL3, WOMAC 0–96].

Health-related quality of life at short-, medium- and long-term follow-up

- Arthroscopic debridement (including some meniscal excision) versus physiotherapy:⁸⁸
 - Short-term Knee Specific Pain Scale (KSPS) (SF-36, physical component) (2 years): MD = −0.20 (−3.53 to 3.13), scale 0–100. Very low certainty evidence; 1 RCT in 168 participants.

Adverse events and serious adverse events

One RCT comparing arthroscopic debridement and HA injection in 120 participants reported no AEs in either arm at 6 months' follow-up.⁹² One NRS comparing prior arthroscopy (unspecified, but not APM) versus no prior arthroscopy also reported no AEs.⁹⁴

Joint replacement (2 years' or more follow-up)

No RCTs reported the incidence of TKR at two or more years' follow-up (although one study reported this outcome at 1-year follow-up).⁸²

Reoperation rate or revision of initial surgical procedure

One RCT in 38 participants reported reoperation rate: 5 of the 19 participants who received HA injection subsequently received arthroscopic surgery during follow-up; no participants had reoperations in the surgery group.⁸²

- Arthroscopic lavage versus IA-HA injection:⁸²
 - Short-term (12 months): Peto's OR = 0.11 (0.02 to 0.68), favouring surgery. Very low certainty evidence.

Type 2: comparisons of one non-regenerative surgical procedure versus another non-regenerative surgical procedure to directly treat early knee osteoarthritis.

Three RCTs in 346 participants reported comparisons of one type of non-regenerative surgical procedure with another for direct treatment of early OA.^{21,85,86} One RCT was a three-arm trial of arthroscopic lavage, arthroscopic debridement and sham surgery.²¹ No studies reported reoperation rate, incidence of subsequent TKR, progression of OA or any PREMs outcomes.

Pain at short- and long-term follow-up

- Arthroscopic debridement versus sham surgery (part of three-arm trial):²¹
 - Short-term (2 years): SMD = −0.01 (−0.39 to 0.37). Very low certainty evidence; 1 RCT in 108 participants at 2 years. [MD = −0.20 (−9.05 to 8.65), Knee Specific Pain Scale (KSPS) (0–100); control group means: baseline 65.0, postoperative 51.6].
- Arthroscopic lavage versus sham surgery (part of three-arm trial):²¹
 - Short-term (2 years): SMD = 0.09 (−0.29 to 0.46). Very low certainty evidence; 1 RCT in 110 participants. [MD = 2.10 (−6.76 to 10.96), KSPS 0–100; control group means: baseline 65.0, postoperative 51.6].
- Arthroscopic lavage (three litres) versus arthroscopic minimal lavage (250 ml):⁸⁶
 - Short-term (12 months): SMD = −0.87 (−1.30 to −0.43), favouring arthroscopic lavage. Low certainty evidence; 1 RCT in 90 participants. [MD = −1.00 (−1.49 to −0.51), VAS 0–10; control group means: baseline 3.6, postoperative 3.3].

- Arthroscopic debridement versus arthroscopic lavage (part of three-arm trial):
 - Short-term (2 years): pain as a continuous outcome – SMD = -0.10 (-0.47 to 0.28). Very low certainty evidence; 1 RCT in 108 participants.²¹ [MD -2.30 (-11.15 to 6.55), KSPS 0–100; control group means: baseline 63.7, postoperative 53.7].
 - Short- and long-term follow-up: pain as a dichotomous outcome 1 RCT in 76 participants reported pain as a dichotomous outcome.⁸⁵ Short-term (12 months): number who were not pain-free at 12 months: RR = 0.23 (0.12 to 0.44), favouring debridement. Very low certainty evidence; 1 RCT in 76 participants, with 39 events (8 and 31) [control group risk 31/36 (86%)]. Long-term (5 years): Number who were not pain-free at 5 years: RR = 0.46 (0.30 to 0.71), favouring debridement. Very low certainty evidence; 1 RCT in 58 participants, with 36 events (13 and 23).

Physical function at short-term follow-up

- Arthroscopic debridement versus sham surgery (part of three-arm trial):²¹
 - Short-term (2 years): SMD = 0.09 (-0.29 to 0.47). Very low certainty evidence; 1 RCT in 108 participants. [MD = 2.60 (-8.15 to 13.35), AIMS2 walking/bending subscale (0–100); control group means: baseline 58.5, postoperative 53.8].
- Arthroscopic lavage versus sham surgery:²¹
 - Short-term (2 years): SMD = -0.10 (-0.47 to 0.28). Very low certainty evidence; 1 RCT in 110 participants. [MD = -2.70 (-13.13 to 7.73), AIMS2 walking/bending subscale (0–100); control group means: baseline 58.5, postoperative 53.8].
- Arthroscopic lavage (3 l) versus arthroscopic minimal lavage (250 ml):⁸⁶
 - Short-term (12 months): SMD = -0.23 (-0.65 to 0.19). Very low certainty evidence; 1 RCT in 90 participants. [MD = -2.00 (-5.59 to 1.59), WOMAC function (0–100); control group means: baseline 28.0, postoperative 21.7].
- Arthroscopic debridement versus arthroscopic lavage (part of three-arm trial):²¹
 - Short-term (2 years): SMD = 0.18 (-0.19 to 0.56). Very low certainty evidence; 1 RCT in 109 participants. [MD = 5.30 (-5.54 to 16.14); AIMS2 walking/bending subscale scale (0–100); control group means: baseline 55.5, postoperative 51.1].

Pain and function at short-, medium- and long-term follow-up

- Arthroscopic lavage (3 l) versus arthroscopic minimal lavage (250 ml):⁸⁶
 - Short-term (12 months): MD = -4.00 (-7.72 to -0.28), favouring arthroscopic lavage; WOMAC (0–100 assumed). LOW certainty evidence; 1 RCT in 90 participants.

Health-related quality of life at short-term follow-up

- Arthroscopic lavage versus sham surgery:²¹
 - Short-term (2 years): MD = 1.10 (-9.14 to 11.34); SF-36 physical function (0–100). Very low certainty evidence; 1 RCT in 106 participants.
- Arthroscopic debridement versus sham surgery:²¹
 - Short-term (2 years): MD = -1.90 (-12.04 to 8.24); SF-36 physical function (0–100). Very low certainty evidence; 1 RCT in 111 participants.
- Arthroscopic debridement versus arthroscopic lavage:²¹
 - Short-term (2 years): MD = 3.00 (-7.12 to 13.12); SF-36 physical function (0–100). Very low certainty evidence; 1 RCT in 108 participants.

Adverse events and serious adverse events

One three-arm RCT in 180 participants, comparing debridement, lavage and sham surgery reported no deaths and two minor postoperative complications, but the group was not stated.²¹ One RCT comparing debridement and lavage in 76 participants reported no deaths in either arm up to 5 years.⁸⁵

Type 3: comparisons of non-regenerative surgical procedures with other joint-preserving or joint replacement surgical options to directly treat early knee OA.

No studies met the inclusion criteria for these comparisons.

Indirect treatment of early osteoarthritis

Type 1: comparisons of non-regenerative surgical procedures versus non-operative management to indirectly treat early knee OA.

Eight RCTs in 1369 randomised participants reported the effect of non-regenerative surgical procedures in comparison with non-operative management to treat specific knee conditions for people in whom some or all participants have early OA.^{83,84,87,89,90,95-97} No studies reported any PREMs outcomes.

Pain at short-, medium- and long-term follow-up

Pain was reported only for one comparison:

- APM versus physiotherapy:
 - Short-term (6 months to 2 years): meta-analysis of 6 RCTs in 1008 participants.^{83,84,87,89,95,97} SMD (random effects) = -0.19 (-0.32 to -0.07), favouring surgery. Heterogeneity: $I^2 = 0\%$, $p = 0.69$. Low certainty evidence.
 - Medium-term (> 2 years to < 5 years): meta-analysis of 2 RCTs in 438 participants.^{83,87} SMD = -0.20 (-0.39 to -0.01). Heterogeneity: $I^2 = 0\%$, $p = 0.60$. Low certainty evidence.
 - Long-term (5 years): meta-analysis of 3 RCTs in 531 participants: SMD (random effects) = -0.01 (-0.32 to 0.30). Heterogeneity $I^2 = 63\%$, $p = 0.07$. Very low certainty evidence.^{83,84,87}

For the APM versus physiotherapy comparison two RCTs also reported pain as a dichotomous outcome (i.e. the number of participants without improvement).^{83,97}

- Short-term (2 years): number with persistent pain at 2 years: RR = 0.62 (0.16–2.47) for a control group value of 96 per 1000. Very low certainty evidence; 1 RCT in 102 participants.⁹⁷
- Medium-term (3 years): number of participants who did not achieve an improvement of more than 10 KOOS points (calculated from the number with an improvement of more than 10 KOOS points): RR = 0.72 (0.37–1.41), for a control group value of 268 per 1000. Very low certainty evidence. One RCT in 118 participants.⁸³

Physical function at short-, medium- and long-term follow-up

- APM versus physiotherapy:^{83,84,87,89,95}
 - Short-term (6 months to 2 years): meta-analysis of five RCTs in 955 participants: SMD (random effects) = -0.13 (-0.26 to 0.00). Heterogeneity: $I^2 = 0\%$, $p = 0.53$. Low certainty evidence.
 - Medium-term (3 years): SMD = -0.29 (-0.65 to 0.08). Very low certainty evidence; 1 RCT in 117 participants.⁸³
 - Long-term (5 years): 3 RCTs in 311 participants data; 1 RCT only reported as-treated analysis, we have reported this separately (but treat with caution).^{83,84,89}
 - Meta-analysis of 2 RCTs in 211 participants: SMD (random effects) = 0.08 (-0.25 to 0.41). Heterogeneity: $I^2 = 32\%$, $p = 0.23$, point estimates show inconsistency. Very low certainty evidence.^{84,89}
 - SMD (as-treated) = 0.22 (-0.19 to 0.63). Very low certainty evidence; 1 RCT in 100 participants.⁸³

Pain and function at short-, medium- and long-term follow-up

- APM versus physiotherapy:^{84,97}
 - Short-term (2 years): meta-analysis of 2 RCTs in 192 participants: SMD (random effects) = 0.10 (–0.19 to 0.38). No heterogeneity: $I^2 = 0\%$, $p = 0.93$. Very low certainty evidence.^{84,97}
 - Long-term (5 years): SMD = 0.26 (–0.15 to 0.68). Very low certainty evidence; 1 RCT in 92 participants.⁸⁴
- APM debridement versus IA steroid injection:⁹⁶
 - Short-term (12 months): SMD = –0.38 (–0.78 to 0.02). Very low certainty evidence; 1 RCT in 98 participants. Two RCTs also reported pain as a dichotomous outcome (i.e. the number of participants without improvement).
- APM versus physiotherapy:⁸⁹
 - Short-term follow-up (2 years): number with a KOOS4 change from baseline not more than 10.1 points: RR = 0.97 (0.47–1.99). Very low certainty evidence; 1 RCT in 126 participants.
- Arthroscopy (mainly APM) versus physiotherapy plus NSAIDs:⁹⁰
 - Short-term follow-up (2 years): number with change from baseline HSS knee points not more than 10: RR = 0.60 (0.32–1.12). Very low certainty evidence; 1 RCT in 73 participants.
 - Medium-term follow-up (3 years): RR = 0.62 (0.36–1.07). Very low certainty evidence; 1 RCT in 73 participants.

Health-related quality of life at short-, medium- and long-term follow-up

- APM versus physiotherapy:^{83,84,87,89,95}
 - Short-term (6 months to 2 years): meta-analysis of 5 RCTs in 955 participants: SMD (random effects) = 0.10 (–0.06 to 0.25), favouring surgery. Heterogeneity: $I^2 = 27\%$, $p = 0.24$, some heterogeneity in the point estimates. Low certainty evidence.
 - Medium-term (3 years): SMD = 0.47 (0.10–0.85). Very low certainty evidence; 1 RCT in 113 participants.⁸³
 - Long-term (5 years): meta-analysis of 2 RCTs in 211 participants: SMD (random effects) = 0.15 (–0.12 to 0.42). Slight heterogeneity in point estimates. $I^2 = 0\%$, $p = 0.34$. Very low certainty evidence.^{84,89}

Adverse events and serious adverse events at short-term and medium-term follow-up

Three RCTs in 811 participants reported AEs. One with no events and two with some events.^{83,87,95} Two studies reported serious adverse events (SAEs) and one RCT reported mortality.^{87,90,95} There is insufficient information to draw conclusions.

- APM versus physiotherapy:
 - AEs – short-term: meta-analysis of 2 RCTs in 672 participants, 41 with events: RR (random effects) = 1.41 (0.77–2.59). Heterogeneity: $I^2 = 0\%$, $p = 0.33$. Very low certainty evidence.^{87,95}
 - SAEs: meta-analysis of 2 RCTs in 669 participants, 12 with events: RR (random effects) = 1.01 (0.33–3.16). Heterogeneity: $I^2 = 0\%$, $p = 0.56$, but there was some variability in the point estimates. Very low certainty evidence.^{87,95}
- Arthroscopy (mainly APM) versus physiotherapy and NSAIDs:
 - Mortality – medium-term (up to 3 years): RR: 2.50 (0.51–12.14). Very low certainty evidence; 1 RCT in 80 participants, 7 with events.⁹⁰

Joint replacement (2 years or more follow-up)

- APM versus physiotherapy:
 - Short-term (2 years): RR = 0.68 (0.12–3.87). Very low certainty evidence; 1 RCT in 321 participants, 5 with events (TKRs).⁹⁵
 - Long-term (5 years):
 - HR = 2.0 (0.84–4.8). Very low certainty evidence; 1 RCT in 351 participants, 25 with events.⁸⁷
 - RR = 3.0 (0.12–75). Very low certainty evidence; 1 RCT in 140 participants, with 1 TKR in the APM arm and none in the control arm.⁸⁹

Regarding joint replacement as an outcome, the two NRS reported adjusted time-to-event data at up to 9 years follow-up.⁹⁹⁻¹⁰¹ The studies were different analyses of the osteoarthritis initiative (OAI) database, a prospective cohort study investigating progression of OA and the incidence of TKR. In the Zikria 2017 study, the presence/absence of a prior arthroscopy treatment was recorded at baseline and in the Rongen 2018 study the intervention was APM during follow-up (with impact on later knee replacement). The OAI population comprised a group of people with OA (KL $\geq$ 2) and a second group including participants at high risk of OA.

- Zikria *et al.* compared participants who had prior APM without preceding knee injury versus no prior APM up to 9 years follow-up. The analysis was adjusted for age, sex, BMI, Physical Activity Scale for the Elderly, WOMAC score and KL grade.
 - HR (adjusted) = 2.10 (1.50–2.92). Low certainty evidence; 1 NRS in 4234 participants, 341 with events (knee replacements). [unadjusted HR = 4.83 (3.52–6.63)].¹⁰⁰
- Rongen *et al.* compared participants who received APM during the follow-up period with a matched group who had no APM in the follow-up period, up to 9 years' follow-up. Propensity scores were used to match the participants 1 : 1 (based on baseline values for 18 covariates). The analysis treated APM as a time-dependent covariate.⁹⁹
 - HR = 3.0 (1.7–5.3), favouring non-operative treatment. Low certainty evidence; 1 NRS in 670 participants, 101 with events (knee replacements).⁹⁹
 - Subgroup analysis results: OA group (KL $\geq$ 2): HR = 2.2 (1.2–3.8); and risk of OA group: HR = 4.7 (2.1–11.1).⁹⁹

Reoperation rate or revision of initial surgical procedure

We conducted two types of meta-analysis, one representing reoperation rate as repeated therapy in the non-operative group or revision in people receiving surgery, and the other including participants switching from the non-operative group to the APM group at any time during follow-up (cross-overs).

- APM versus physiotherapy short-term follow-up (2 years):^{83,84,87,89,95}
 - Excluding cross-overs: meta-analysis of five RCTs in 1058 participants (1 with no events), 13 with events: RR = 3.9 (1.1–13.8), favouring non-operative management. No heterogeneity $I^2 = 0\%$, $p = 0.96$. Very low certainty evidence.
 - Including cross-overs: meta-analysis of 5 RCTs in 1058 participants, 158 with events: RR 0.13 (0.05–0.32), favouring surgery. Heterogeneity $I^2 = 55\%$, $p = 0.06$ (but all results in the same direction). Very low certainty evidence.

Progression of osteoarthritis (2 years' follow-up or more)

- APM versus physiotherapy:
 - Short-term (2 years): mean (SD) KL grade (calculated from numbers with each KL grade): SMD = –0.08 (–0.31 to 0.15). Very low certainty evidence; 1 RCT in 102 participants.⁹⁷
 - Long-term (5 years), mean KL grade: SMD = –0.01 (–0.28 to 0.26). Very low certainty evidence; 1 RCT in 129 participants.⁸⁹
- APM versus physiotherapy:
 - Short-term and long-term follow-up (2–5 years): meta-analysis of 3 RCTs in 310 participants, 28 with events: RR = 0.98 (0.49–1.94). No heterogeneity $I^2 = 0\%$, $p = 0.92$. Very low certainty evidence.^{84,89,97}
- Prior APM without preceding knee injury versus no prior APM:
 - Long-term (8 years): evolution of JSN (for participants with baseline JSN OARSI grade < 3):
 - Adjusted HR = 1.27 (1.00–1.63) (adjusted for age, sex, BMI, Physical Activity Scale for the Elderly, WOMAC score, KL grade), that is more progression for the prior surgery group. Low certainty evidence; 1 NRS in 3973 participants, 1187 with events [unadjusted HR = 2.04 (1.60–2.60)].^{100,101}
 - Subgroup, KL $\geq$ 2: HR (adjusted for all but KL grade) = 1.45 (1.13–1.87); all KL grades: HR = 1.80 (1.41–2.30).

Type 2: comparisons of one non-regenerative surgical procedure versus another non-regenerative surgical procedure to indirectly treat early knee OA.

Four RCTs in 440 randomised participants reported comparisons of one type of non-regenerative surgical procedure versus another to treat specific knee conditions for people in whom some or all participants have early OA.^{22,80,91,93} Of these studies reported APM versus sham surgery in people with a meniscal tear.^{22,91} Sham surgery was described as skin-incisions-only by 1, smaller study (34 participants) which also manipulated the knee; the other, larger study (144 participants) described sham surgery as following a standard APM procedure, which may have given a closer simulation.^{22,91} Two studies investigated combination surgery in people who had both a meniscal tear and chondral defects: 1 study investigated the addition of chondroplasty to APM versus no chondroplasty in 161 participants with meniscal tear and unstable chondral lesions of Outerbridge grade II, III or IV in any compartment of the knee and the other study reported on differences between APM plus radiofrequency chondroplasty and APM plus mechanical shaver debridement (MSD) in 60 participants with a medial meniscus tear and idiopathic Grade III cartilage defect(s) of the medial femoral condyle.^{80,93} Only one study reported reoperation rate, incidence of TKR and radiological progression of OA at medium- and long-term follow-up.⁹³ No studies reported progression of OA or any PREMs outcomes. An asterisk * indicates we converted the results to a negative scale.

Pain at short-term follow-up

- APM versus sham surgery:
 - Short-term (2 years): meta-analysis of 2 RCTs in 178 participants. SMD (random effects) = -0.29 (-0.99 to 0.40). Heterogeneity: $I^2 = 70\%$, $p = 0.07$. VERY LOW certainty evidence. MD for Roos *et al.* 2018 = -17.50 (-33.84 to -1.16), 34 participants; KOOS pain (0–100); control group means: baseline $\times$ 54.1, postoperative $\times$ 38.7. MD for Sihvonen *et al.*²² = 0.00 (-0.83 to 0.83), 144 participants; VAS pain after exercise (0–10); control group means: baseline 6.1, postoperative 2.3.^{22,91}
- APM plus chondroplasty versus APM (in people with meniscal tears and chondral defects):
 - Short-term (12 months): MD = -0.60 (-1.21 to 0.01), VAS pain (0–10), favouring APM plus chondroplasty. Low certainty evidence; 1 RCT in 161 participants. (control group means: baseline 4.5, postoperative 2.0).⁸⁰
- APM + arthroscopic radiofrequency chondroplasty (RFC) versus APM + MSD (in people with meniscal tears and chondral defects):
 - Short-term (12 months): MD = -23.30 (-28.94 to -17.66), KOOS pain (0–100), favouring RFC. Low/very low certainty evidence; 1 RCT in 60 participants.⁹³ [Control group means: baseline $\times$ 90.4 (i.e. very high pain), postoperative $\times$ 42.1.]
 - Medium-term (4 years): MD = -19.40 (-30.21 to -8.59), favouring RFC; KOOS pain scale (0–100). Very low certainty evidence; 1 RCT in 40 participants.⁹³ (Control group means: baseline $\times$ 87.9, postoperative $\times$ 55.7.) Baseline difference of 4.7 points.
 - Long-term (10 years): MD -4.50 (-15.83 to 6.83); KOOS pain (0–100). Very low certainty evidence; 1 RCT in 22 participants (much missing data) (control group means: baseline $\times$ 88.3, postoperative $\times$ 59.3).⁹³

Physical function at short-term follow-up

- APM versus sham surgery:
 - Short-term (2 years): MD = -13.40 (-27.74 to 0.94), KOOS activities of daily living (ADL) scale 0–100. Very low certainty evidence; 1 RCT in 34 participants.⁹¹
- APM plus chondroplasty versus APM:
 - Short-term (12 months): MD = -1.90 (-6.87 to 3.07), WOMAC function (0–100), for a control group score of 26.6 points. Low certainty evidence; 1 RCT in 161 participants.⁸⁰
- APM plus RFC versus APM plus MSD:
 - Short-term (12 months): MD = -22.70 (-26.46 to -18.94), KOOS ADL (0–100), favouring RFC. Low/very low certainty evidence; 1 RCT in 60 participants.⁹³

- Medium-term (4 years): MD = -19.00 (-29.00 to -9.00), KOOS ADL (0–100), favouring RFC. Very low certainty evidence; 1 RCT in 40 participants⁹³ (control group baseline 11.4, postoperative score 50.9).⁹³
- Long-term (10 years): MD -29.70 (-41.05 to -18.35), KOOS ADL (0–100), favouring RFC. Very low certainty evidence; 1 RCT in 22 participants [control group baseline 10.1, postoperative 27.2]. Baseline difference 7.1 points.⁹³

Pain and function at short-term follow-up

Two RCTs in 178 participants reported this outcome for the comparison of APM versus sham surgery, but Sihvonon *et al.* reported both final values and change scores for the Lysholm scale and Roos *et al.* reported change scores for the KOOS score.^{22,91} For this outcome, we report the final values results separately and meta-analyse the change from baseline results using the SMD. A negative mean difference or SMD indicates less pain and function for the APM group.

- APM versus sham surgery:
 - Short-term (2 years): final values: MD = 2.60 (-2.03 to 7.23). Very low certainty evidence; 1 RCT in 144 participants. Change scores: meta-analysis of 2 RCTs in 178 participants: SMD (change scores) = -0.02 (-0.56 to 0.52). Heterogeneity: $I^2 = 55\%$, $p = 0.14$. Very low certainty evidence.^{22,91}
- APM plus RFC versus APM plus MSD: 1 RCT in 60 participants compared APM plus RFC versus APM plus MSD.⁹³
 - Short-term (12 months): MD = -23.90 (-27.93 to -19.87), KOOS total (0–100), favouring RFC. Low/very low certainty evidence; 1 RCT in 60 participants.
 - Medium-term (4 years): MD = -18.60 (-28.15 to -9.05), KOOS total (0–100), favouring RFC. Very low certainty evidence; 1 RCT in 40 participants.
 - Long-term (10 years): MD -17.70 (-25.20 to -10.20), KOOS total (0–100), favouring RFC. Baseline difference of 8.7 points very low certainty evidence; 1 RCT in 22 participants.

Health-related quality of life at short-term follow-up

- APM versus sham surgery:
 - Short-term (2 years): meta-analysis of 2 RCTs in 178 participants: SMD (random effects) = 0.25 (-0.37 to 0.87). Heterogeneity: $I^2 = 64\%$, $p = 0.10$. Very low certainty evidence.^{22,91}
- APM plus chondroplasty versus APM:
 - Short-term (12 months): MD = -0.06 (-0.37 to 0.25), SF-36 (0–100). Low certainty evidence; 1 RCT in 161 participants.⁸⁰
- APM plus RFC versus APM plus MSD.⁹³
 - Short-term (12 months): MD = 24.00 (16.78–31.22), KOOS QoL (0–100), favouring RFC. Low/very low certainty evidence; 1 RCT in 60 participants (control group baseline 10.8, postoperative score 56.2).
 - Medium-term (4 years): MD = 14.10 (2.13–26.07), KOOS QoL (0–100), favouring RFC. Very low certainty evidence; 1 RCT in 40 participants (control group baseline 9.5, postoperative score 52.9). Baseline difference of 5 points.
 - Long-term (10 years): MD 8.80 (-4.03 to 21.63), KOOS QoL (0–100). Very low certainty evidence; 1 RCT in 22 participants (control group baseline 6.9, postoperative score 41.6). Baseline difference of 10.9 points.

Adverse events and serious adverse events at short-term follow-up

- APM versus sham surgery:
 - AEs in the short-term (2 years): RR = 0.67 (0.12–3.61). Very low certainty evidence; 1 RCT in 44 participants, 5 with events.⁹¹
 - SAEs in the short-term (2 years): meta-analysis of 2 RCTs in 190 participants, 4 with events: Peto's OR = 8.12 (1.10–59.7). No heterogeneity: $I^2 = 0\%$, $p = 1.00$. Very low certainty evidence.^{22,91}
 - Mortality in the short-term (2 years): 1 RCT in 146 participants, with 1 event. Very low certainty evidence.²²

- APM plus chondroplasty versus APM:
 - AEs in the short-term (12 months): 1 RCT in 190 participants reported no AEs in either arm.⁸⁰
- APM plus RFC versus APM plus MSD:
 - One RCT in 60 participants reported 3 deaths over a 10-year follow-up period. One death occurred in the MSD group by medium-term follow-up (4 years), and the other two, one in each group, by 10-year follow-up.⁹³

Total knee replacement rate at 2 years or longer

- APM versus sham surgery:
 - Short-term follow-up (2 years): 1 RCT in 144 participants, 3 with events. Very low certainty evidence.²²
- APM plus RFC versus APM plus MSD:
 - Medium-term follow-up (4 years): Peto's OR = 0.17 (0.04–0.67). One RCT in 60 participants, 9 with events. Very low certainty evidence.⁹³

Reoperation rate at short-term follow-up

Two RCTs in 190 participants comparing surgery and sham surgery reported reoperation rates at 2 years, and 1 RCT in 60 participants comparing APM plus RFC with APM plus MSD showed a Kaplan–Meier plot for all-cause revision up to 10 years and reported a log-rank *p*-value.^{22,91,93}

- APM versus sham surgery, short-term follow-up (2 years):^{22,91}
 - Excluding switching from sham to APM: meta-analysis of 2 RCTs in 190 participants, 3 with events: RR = 4.0 (0.5–35.6). No heterogeneity $I^2 = 0\%$, $p = 0.80$. Very low certainty evidence.
 - Including switching from sham to APM: meta-analysis of 2 RCTs in 190 participants, with 17 events: RR 0.24 (0.07–0.81), favouring APM. No heterogeneity $I^2 = 0\%$, $p = 0.42$. Very low certainty evidence. Control group risks were 360 per 1000 in the smaller study and 80 per 1000 in the larger study; the latter corresponds to 61 fewer people per 1000 having reoperations in the APM group (from 15 to 74 fewer).
- APM plus RFC versus APM plus MSD:⁹³
 - Long-term follow-up (10 years): HR = 0.36 (0.15–0.86), log-rank *p*-value: 0.022; 7/30 revisions in the RFC group (1 knee replacement, 2 osteotomies, 1 revision arthroscopy and 3 others), and 18/30 revisions in the MSD group (11 knee replacements, 5 osteotomies, 2 revision arthroscopies). This corresponds to 319 fewer people with revisions per 1000 receiving APM plus RFC (from 55 to 472 fewer). Low certainty evidence.

Progression of osteoarthritis at short- medium- and long-term follow-up

One RCT assessed radiological progression with time in the 22 (of 60) participants available at 10 years, measuring the medial compartment joint width at 1, 4 and 10-year follow-up.⁹³ Results are reported for medial JSW, reversed in sign to indicate progression (a negative MD favours RFC): this was similar in the two groups at baseline and narrowed in both groups over time.

- APM plus RFC versus APM plus MSD:⁹³
 - Short-term (12 months): MD = –1.2 (–2.45 to 0.05) mm of joint space; baselines 5.6 mm and 5.4 mm. Very low certainty evidence.
 - Medium-term (4 years): MD = –1.50 (–2.75 to –0.25) mm. Very low certainty evidence.
 - Long-term (10 years): MD = –1.30 (–3.28 to 0.68) mm. Very low certainty evidence.

Type 3: comparisons of non-regenerative surgical procedures with other joint-preserving or joint replacement surgical options to indirectly treat early knee osteoarthritis.

No studies met the inclusion criteria for these comparisons.

Conclusions

There was limited evidence to support the use of non-regenerative surgical interventions to treat early OA of the knee. In considering the evidence from studies where procedures were used to directly treat early knee OA, we found two interventions where evidence was available.

Firstly, there were a number of studies that investigated arthroscopic lavage; however, they provided no convincing evidence that this procedure was effective. This is line with previous NICE treatment guidelines that have recommended: 'Current evidence suggests that arthroscopic knee washout alone should not be used as a treatment for OA because it cannot demonstrate clinically useful benefit in the short or long term'.⁹⁸

There is limited evidence to support arthroscopic debridement procedures used to directly treat early knee OA. We found four RCTs and one NRS investigating this procedure (see [Table 3](#)). Two of the RCTs were underpowered (approximately 30 patients in each) and did not provide data that could be taken forward to the health economic analysis.^{81,82} Two studies provided evidence that could be taken forward to the health economic analysis. There is 1 RCT in 168 participants reporting comparison of arthroscopic debridement (including some meniscal excision) versus physiotherapy, with some benefit in functional outcome seen with the surgical procedure.⁸⁸ This evidence met the criteria to be taken forward for health economic analysis, but it is recognised that the evidence is of low or very low certainty, mainly because studies are at high risk of bias or there may be imprecision in the estimates.

Other studies consisted of a disparate array of comparisons of non-regenerative surgical procedures and non-operative management mainly from small studies, which could not be synthesised together.

In indirect early OA treatment of specific knee conditions, the evidence identified considers two main types of morphological feature: meniscal tears and chondral defects. For the comparison of non-regenerative surgery versus non-operative management, only the comparison APM versus physiotherapy was reported in the included studies. Interpretation of these data was challenging due to the lack of direct comparison, high potential for bias and the low certainty of the evidence produced.

Chapter 4 Non-regenerative surgical procedures used to treat early osteoarthritis in the hip

Results of the search

The search and successive updates up to January 2021 generated 7250 records. After deduplication and first assessment based on title and abstract screening, we obtained 213 in full text, comprising 56 systematic reviews and 157 reports of potential RCTs or comparative NRS ([Figure 2](#)). The search strategy is recorded in [Appendix 1](#).

We included six intervention studies: four RCTs and one comparative NRS.¹⁰²⁻¹⁰⁷ We excluded 120 studies and assessed 56 systematic reviews and overviews, none of which could be used alone to answer the review questions. We excluded 120 studies, of which 4 were RCTs. Five excluded NRS compared hip arthroscopy with non-operative management or no arthroscopy; 8 compared labral reconstruction and labral repair; 9 compared labral reconstruction and labral repair; and 30 had ineligible comparisons of surgical approaches or surgical techniques. One important reason for exclusion in the 120 studies was a difference between groups in the indications for surgery. Other common reasons for exclusion were an ineligible comparison in 32 studies, an ineligible study design in 13 and not taking into account confounding factors in 7 NRS.

Ongoing studies characteristics

We are aware of three ongoing RCTs. One RCT is investigating the efficacy of hip arthroscopic procedures compared to sham surgery (diagnostic arthroscopy only) on PROMs for patients with symptomatic and clinical and radiological findings related to impingement and/or labral tears (HIPARTI Study; NCT02692807). The study is planned to report at 1, 2, 5 and 10 years' follow-up.¹⁰⁸

The second RCT (NCT03909178) is investigating the efficacy of arthroscopic surgical labral repair plus a 24-week course of supervised, core-based physiotherapy compared with physiotherapy treatment alone in people 40 years and over, with tears of the acetabular labrum and mild to moderate OA severity. The study will report at 1, 2, 5 and 10 years, and the 12 months' results have been published since the end date of our search.¹⁰⁹ Briefly, the findings are, for pain and function at 12 months, adjusted for baseline, MD = -12.11 (-20.96 to -3.27) iHOT-33 points (scale 0-100). However, the authors note that 64% of physiotherapy participants switched to the surgical intervention.

A final RCT is investigating the two established surgical treatments, labral repair versus labral debridement for acetabular labral tears, with KL OA severity grades less than 2 (ALERT study; NCT03257709). The study will report at 6 months' and 2 years' follow-up, and the estimated final completion date was December 2019, but there is no record of results being published at the time of this report.¹¹⁰

Included studies characteristics

The main features of the characteristics of included studies are summarised in [Table 5](#).

Study details

A total of five RCTs met the inclusion criteria for this review.¹⁰²⁻¹⁰⁶ One comparative NRS was also included.¹⁰⁷

Characteristics of included studies

A total of 5 RCTs involving 900 participants were included in this review. For these studies, the median sample size was 214 (IQR 80-222); 3 RCTs had more than 200 participants with the largest containing 348.^{103,105,106} Pre-published protocols were available for four of the trials.¹⁰³⁻¹⁰⁶ One comparative NRS conducted from a large database was also included.¹⁰⁷

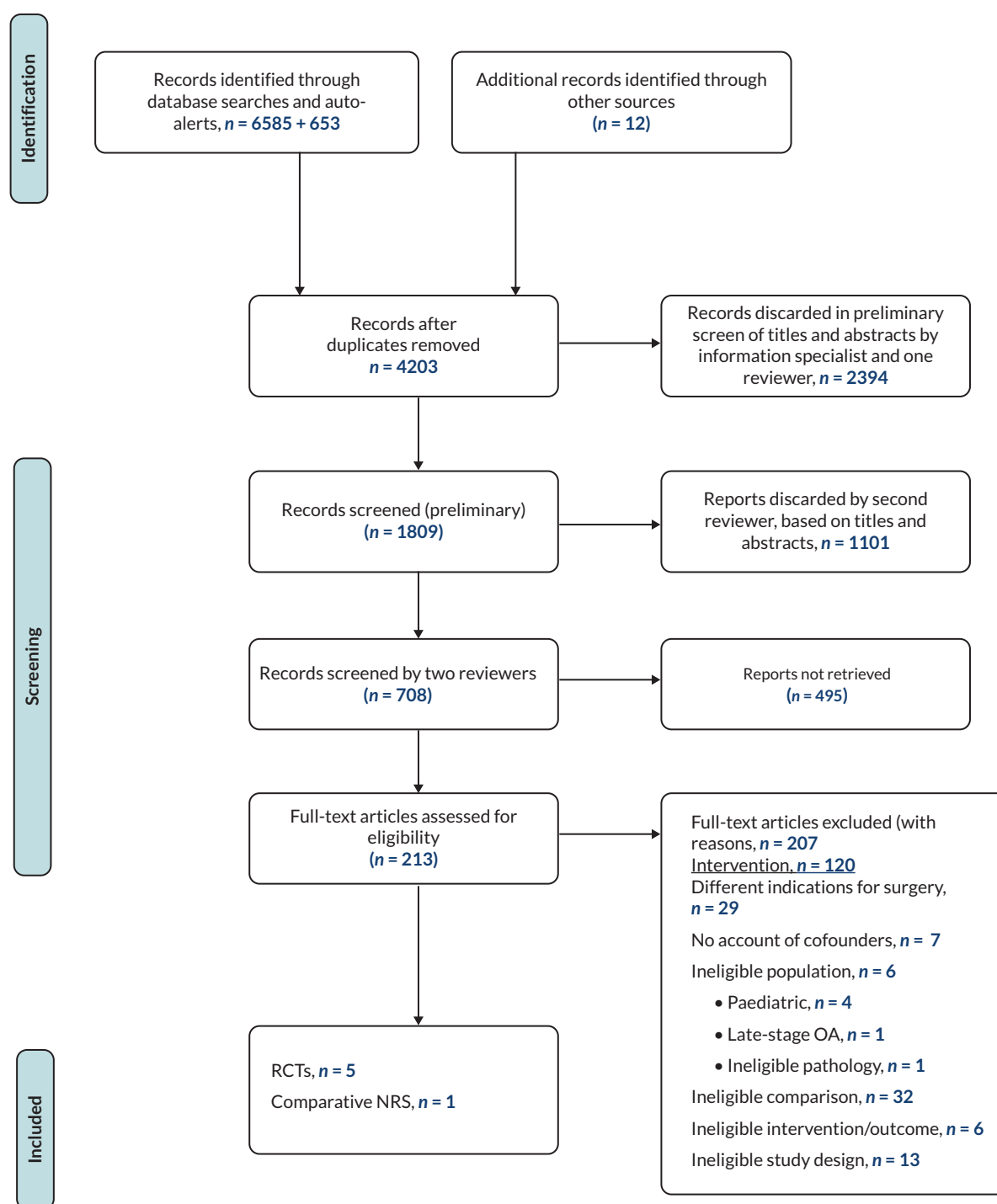


FIGURE 2 PRISMA diagram: non-regenerative surgery for treatment of early hip OA.

Of the five RCTs, two were conducted in the UK,^{103,105} two were conducted in the USA,^{102,104} with one conducted in Canada.¹⁰⁶ Two trials were single-centre studies.^{102,104} The remaining trials were conducted over 7, 10 and 23 sites, respectively.^{103,105,106} All studies randomised interventions to individual participants. The NRS was conducted in the USA and was a retrospective cohort study.¹⁰⁷

Of the six studies included, four received non-commercial funding, one reporting no funding and in one study the source of funding was not stated.

Population, intervention, control/comparison, outcome characteristics

Summary details of the population, intervention, control/comparison, outcome (PICO) characteristics are presented in [Table 5](#).

Surgical interventions assessed

The following surgical interventions assessed in the included studies are listed below. We have used the term 'osteocondroplasty' to describe a combination of arthroscopic procedures of hip where bone or cartilage is debrided, or labrum is debrided or repaired.

- Arthroscopic osteochondroplasty: four RCTs and one NRS.¹⁰³⁻¹⁰⁷
- Arthroscopic labral repair: one RCT.¹⁰²

Comparator interventions assessed

The following non-operative comparators were reported in the included studies :

- Physiotherapy: three RCTs.¹⁰³⁻¹⁰⁵
- Mixed conservative care (physiotherapy, steroid injection and other): one NRS.¹⁰⁷

The following control surgical comparators were reported:

- Arthroscopic labral debridement: one RCT.¹⁰²
- Arthroscopic lavage ($\pm$ labral repair) (as a surgical comparator): one RCT.¹⁰⁶

No studies explicitly reported NICE 'core' guidance as a comparator (combination of information, exercise/physical therapy and weight loss).

Outcomes

All studies reported at least one major outcome described in our protocol. Most outcomes related to combined pain and function, pain or physical function, HRQoL and conversion to THA/revision or reoperation rate. Four of the included RCTs only reported results at short-term follow-up (6 months–2 years).¹⁰³⁻¹⁰⁶ One RCT reported medium-term follow-up.¹⁰² The NRS had short-term follow-up.¹⁰⁷

- Pain and function: four RCTs.¹⁰³⁻¹⁰⁶
- Pain: two RCTs.^{105,106}
- Physical function: four RCTs.^{102,104-106}
- HRQoL: three RCTs.^{103,105,106}
- PREMs: two RCTs.^{102,104}
- Sports activity: three RCTs.^{102,104,105}
- Surgical complications and AEs: four RCTs.¹⁰³⁻¹⁰⁶
- Joint replacement (2 years' follow-up or more): three RCTs.¹⁰³⁻¹⁰⁵
- Reoperation rate/revision of initial surgical procedure, including any new procedure, and time to revision or other reoperation: three RCTs and one NRS.¹⁰⁴⁻¹⁰⁷
- Radiological Outcomes (progression of OA – 2 years' follow-up or more): two RCT.^{104,105}

Results

We present the results in three main sections reflecting the different studies; type 1: comparisons of non-regenerative surgical procedures versus non-operative management or no specified non-operative treatment/no surgery; type 2: comparisons of one non-regenerative surgical procedure versus another non-regenerative procedure and type 3: comparisons of a non-regenerative procedure versus another surgical procedure. Within each category, where available, we report each outcome at short-, medium- and long-term follow-up for each of the individual comparisons, giving results separately for RCTs and NRS.

TABLE 5 Summary of included non-regenerative surgery studies characteristics for treatment of early hip OA

Study ID, design and country	Study size and duration	Pathology and indication for surgery (KL/Tönnis grade)	Population characteristics: mean age (SD), sex; male (%), BMI (kg/m ²)	Intervention	Comparator	Adjustment for confounding factors (method/factors adjusted for)	Outcomes
<i>Type 1: comparisons of non-regenerative surgical procedures in the hip vs. non-operative management (RCTs and NRSs)</i>							
Griffin 2018 ¹⁰³ (FASHION) RCT (23 centres) UK	348 (171 and 177) 12 months	FAI (radiographic features of cam or pincer morphology), hip pain and Tönnis grade ≤ 1	<ul style="list-style-type: none"> Age: 35 (9.7) and 35 (9.4) Sex: 100 (58%) and 113 (64%) BMI: N/R 	Osteochondroplasty (correction of shape abnormalities and labral and cartilage pathology)	Physiotherapy (personalised and progressive conservative care programme)	RCT	Combined pain and function, QoL, AEs/SAEs, reoperation/revision
Mansell 2018 ¹⁰⁴ RCT (1 centre) USA	80 (40 and 40) 2 years	FAI	<ul style="list-style-type: none"> Age: 29.7 (7.4) and 30.6 (7.4) Sex: 21 (52.5%) and 26 (65%) BMI: 28.2 (4.4) and 27.5 (4.3) 	Osteochondroplasty	Physiotherapy	RCT	Combined pain and function, function, PREMs, sports, progression, AEs/SAEs, reoperation/revision
Palmer 2019 ¹⁰⁵ RCT (7 centres) UK	222 (112 and 110) Approximately 6 months	FAI	<ul style="list-style-type: none"> Age: 36.4 (9.6) and 36.0 (9.9) Sex: 38 (34%) and 37 (34%) BMI: 25.9 (4.8) and 26.6 (4.8) 	Osteochondroplasty	Physiotherapy	RCT	Combined pain and function, function, QoL, PREMs, sports, AEs/SAEs, reoperation/revision
Scott 2020 ¹⁰⁷ NRS (Humana claims database) USA	864 (432 and 432) At least 2 years	FAI/labral tears	<ul style="list-style-type: none"> Age: 118 (27%) ≥ 50 years old, and 123 (28.5%) ≥ 50 years old Sex: 147 (34%) and 151 (35%) BMI: N/R 	Arthroscopy	Conservative mixed	Propensity score matching 1 : 1 (age, sex, Charlson Comorbidity Index (CCI), ethnicity, obesity, diabetes and smoking status)	Reoperation/revision
<i>Type 2: non-regenerative procedure vs. another non-regenerative procedure (RCTs only)</i>							
Krych 2013 ¹⁰² RCT (1 centre) USA	36 (18 and 18) Mean 32 months	FAI	<ul style="list-style-type: none"> Age: 38 (range 20–59) and 39 (range 19–55) Sex: 0 (0%) (all women) BMI: N/R 	Labral repair	Labral debridement	RCT	Function, PREMs, sports

continued

TABLE 5 Summary of included non-regenerative surgery studies characteristics for treatment of early hip OA (continued)

Study ID, design and country	Study size and duration	Pathology and indication for surgery (KL/Tönnis grade)	Population characteristics: mean age (SD), sex; male (%), BMI (kg/m ²)	Intervention	Comparator	Adjustment for confounding factors (method/factors adjusted for)	Outcomes
Ayeni 2020 ¹⁰⁶ (FIRST) RCT (10 centres) Canada, Finland, Denmark	214 (108 and 106)	FAI	<ul style="list-style-type: none">• Age: 36.7 (8.2) and 35.4 (8.8)• Sex: 66 (61.1%) and 67 (63.2%)• BMI NR	Osteochondroplasty ± labral repair	Arthroscopic lavage	RCT	Combined pain and function, pain, function, QoL, AEs/SAEs, reoperation/revision
Type 3: non-regenerative surgical procedure vs. joint-preserving surgical procedures (RCTs only)							
There were no studies found in this section							

We report the mean difference (MD) where there are common outcome measurement scales and the SMD in all other cases, with 95% CIs. A negative SMD value means that the effect favours surgery, except for HRQoL (in which case a negative value favours the non-operative/comparator treatment). Where appropriate, random effects meta-analysis was performed.

The studies' findings are summarised in [Table 6](#). Risk-of-bias assessment for included studies is presented in [Appendix 2](#).

Type 1: comparison of non-regenerative surgical procedures versus non-operative management

Three RCTs in 650 participants and 1 NRS in 864 participants compared hip arthroscopy with non-operative management. All the studies had only short-term follow-up data (6 months–2 years), and we did not identify any evidence on the effect of interventions at medium-term and longer-term follow-up. We note that one of the RCTs (and one ongoing RCT) plan to report at medium-term follow-up times.

Pain at short-term: very low certainty evidence from 1 RCT in 179 participants suggests a moderate-sized benefit favouring osteochondroplasty over physiotherapy of 11.1 HAGOS pain scale points (0–100), 95% CI 4.7 to 17.5. Both treatments lead to an improvement in pain compared with baseline values. For example, the improvement over baseline for the physiotherapy group was 10 pain points. Similarly, for combined pain and function at short term: low certainty evidence from 3 RCTs in 574 participants suggests there may be a moderate-sized benefit of 11.2 iHOT-33 points (scale 0–100), 95% CI 5.4 to 17.1, favouring osteochondroplasty. Both treatments improve pain and function outcomes compared to baseline (e.g. 9.8 points for physiotherapy).

Regarding QoL at short term, there is low certainty evidence from 2 RCTs in 468 participants suggests no important difference between non-regenerative surgical and conservative treatments. [MD = 3.69 EQ-5D VAS points (0–100), 95% CI –0.02 to 7.40]. Evidence for the outcomes of both physical function and function during sports at short term is of very low certainty, displaying high levels of heterogeneity which may have resulted from 70% of participants switching from the conservative to the surgical intervention in one of the studies.

Very low certainty evidence from 1 large NRS in 846 participants suggests a small increase in the number of people having a THA at 2 years' follow-up in the surgical group when compared to conservative care, with an absolute risk difference of 14 more THAs per 1000 (14 fewer to 60 more), for a conservative group risk of 53 per 1000. Very low certainty evidence from 2 RCTs in 302 participants suggests that, in the context of a trial, those receiving surgery may be less likely to have revision surgery than those receiving physiotherapy are likely to switch to the surgical intervention: RR 0.05, 95% CI 0.01 to 0.23.

The RCT data for arthroscopic osteochondroplasty versus physiotherapy in more detail:

Pain at short-term follow-up one RCT¹⁰⁴

- Short-term (8 months): MD = –11.10 (–17.53 to –4.67); that is moderate benefit favouring surgery. HAGOS score, pain subscale (0–100). Very low certainty evidence: 1 RCT in 180 participants.

Pain and function at short-term follow-up (6 months–2 years) three RCTs^{103–105}

- Short-term: meta-analysis of 3 RCTs in 574 participants MD (random effects) = –11.23 (–17.05 to –5.40), favouring surgery; iHOT-33 scale (0–100). Heterogeneity: $I^2 = 37\%$. One study reported 70% of participants switched from the physiotherapy group to the arthroscopy group (SA removing this study increased the heterogeneity to 60%). Low certainty evidence.

Physical function at short-term follow-up two randomised controlled trials^{104,105}

- Short-term: meta-analysis of 2 RCTs in 262 participants: MD (random effects) = –3.28 (–15.97 to 9.41); HOS ADL subscale (0–100). Heterogeneity: $I^2 = 82\%$. One study reported 70% of participants switched from the physiotherapy

TABLE 6 Summary of findings table for non-regenerative surgery studies for treatment of early hip OA

Non-regenerative surgical procedures compared with non-operative management for FAI or early OA of the hip						
Patients or population: people with FAI or early OA of the hip Intervention: non-regenerative surgical procedure			Settings: primary and secondary care Comparison: non-operative management or no non-regenerative surgery			
Continuous outcomes	Score with non-operative management	MD (back-transformed)	SMD (95% CI) findings	Sample size (no. studies)	Evidence certainty (GRADE)	Comments
RCT evidence: arthroscopic osteochondroplasty vs. non-operative management (physiotherapy). NRS evidence: arthroscopic osteochondroplasty vs. non-operative management (physiotherapy, medication, steroid injection and other)						
Pain at 8 months (short-term) <i>Palmer 2019</i> ¹⁰⁵	HAGOS score pain subscale (0–100) = 37.9 (baseline 47.4)	MD: –11.10 (–17.53 to –4.67), that is a moderate benefit, favouring surgery (complete case)	–0.50 (–0.80 to –0.20) that is a moderate benefit favouring surgery (complete case)	179 (1 RCT)	⊕⊕⊕⊕ Very low ¹	RCTs. Pain analysed as a negative outcome. Negative MD favours surgery. Improvement in HAGOS pain score over baseline for non-operative management group = 9.5
Pain at medium- and long-term duration – not measured						
Pain and function at short-term (6 months–2 years) <i>Griffin 2018</i> ¹⁰³ <i>Mansell 2018</i> ¹⁰⁴ <i>Palmer 2019</i> ¹⁰⁵	Median iHOT-33 score (0–100) = 54.5 Median baseline score = 64.4	Meta-analysis: MD: –11.23 (–17.05 to –5.40) that is a moderate benefit, favouring surgery	SMD meta-analysis: –0.42 (–0.63 to –0.21) that is a moderate benefit favouring surgery	574 (3 RCTs)	⊕⊕⊕⊕ Low ²	RCTs. Pain and function analysed as a negative outcome. Negative SMD and MD favour surgery. All studies used iHot-33 score, so meta-analysis as the mean difference. Some heterogeneity ($I^2 = 37\%$), but all findings in the same direction. Improvement in iHOT-33 score over baseline for the non-operative management group, difference in medians = 9.9
Pain and function at medium- and long-term duration – not reported in any study						
Function at short-term (6 months–2 years) <i>Mansell 2018</i> ¹⁰⁴ <i>Palmer 2019</i> ¹⁰⁵	Mean HOS ADL score (0–100) = 28.9* (Mean baseline score = 34.9*)	Meta-analysis (random effects): MD: –3.28 (–15.97 to 9.41) (complete case) (Palmer 2019: –9.20 (–14.77 to –3.63) – without switching)	SMD: –0.17 (–0.81 to 0.47)	262 (2 RCTs)	⊕⊕⊕⊕ Very low ³	RCT. Function analysed as a negative outcome. Negative MD favours surgery. Both studies used HOS ADL so meta-analysis as MD. High inconsistency ($I^2 = 82\%$). MD Mansell: + 3.80 (–5.63 to 13.23) Palmer: Palmer 2019: –9.20 (–14.77 to –3.63) 70% switching rate from physiotherapy to osteochondroplasty in Mansell 2018, which had a complete case result of + 3.80 (–5.63 to 13.23) and an as-treated analysis result of –11.90 (–25.81 to 2.01). Additionally, this study may not have been fully representative (some indirectness) of the FAI population – military personnel, but contributes < 50% of the weight Improvement in HOS ADL score over baseline for the non-operative management group, difference in means = 6

Non-regenerative surgical procedures compared with non-operative management for FAI or early OA of the hip						
Patients or population: people with FAI or early OA of the hip Intervention: non-regenerative surgical procedure			Settings: primary and secondary care Comparison: non-operative management or no non-regenerative surgery			
Continuous outcomes	Score with non-operative management	MD (back-transformed)	SMD (95% CI) findings	Sample size (no. studies)	Evidence certainty (GRADE)	Comments
QoL, 6 months–2 years (short-term) <i>Griffin 2018</i> ¹⁰³ <i>Palmer 2019</i> ¹⁰⁵	Mean EQ-5D VAS score (0–100) = 68.7 (Mean baseline score = 67.1)	Meta-analysis: MD: 3.69 (–0.02 to 7.40) that is a small benefit, favouring surgery	Meta-analysis: 0.18 (–0.00 to 0.36) that is small benefit, favouring surgery	468 (2 RCTs)	⊕⊕⊕⊕ Low ²	RCT. QoL analysed as a positive outcome. Both studies used EQ-5D (0–100) so we meta-analysed as MD; higher = better No inconsistency Positive MD favours surgery. Improvement in EQ-5D VAS (0–100) score over baseline for the conservative group, difference in means = 1.6
Binary outcomes	Absolute risk non-operative	Absolute risk difference (95% CI)	Relative effect (95% CI)	Number of participants (studies)	Evidence Certainty (GRADE)	Comments
Total hip replacement at short-term (minimum 2 years' follow-up) <i>Scott 2020</i> ¹⁰⁷	53 people per 1000 had a THA	14 more people had a THA per 1000 (from 14 fewer to 60 more)	RR = 1.26 (0.74 to 2.14)	864 (1 NRS) 52 people with THA events	⊕⊕⊕⊕ Very low ⁴	NRS. Propensity score matching Negative outcome, RR < 1 favours surgery Kaplan–Meier plot shows early benefit for surgery group up to 18 months
Reoperation rate short-term (6 months–2 years) <i>Mansell 2018</i> ¹⁰⁴ <i>Palmer 2019</i> ¹⁰⁵	Low risk: 55 people per 1000 had reoperation (<i>Palmer 2019</i>) High risk: 700 people per 1000 had reoperation (<i>Mansell 2018</i>)	52 fewer people had reoperations per 1000 (from 54 fewer to 42 fewer) 665 fewer people had reoperations per 1000 (from 693 fewer to 539 fewer)	RR 0.05 (0.01 to 0.23)	302, 35 with events (2 RCTs)	⊕⊕⊕⊕ Very low ¹	RCT. Negative outcome, RR < 1 favours surgery. No inconsistency
SAEs at short-term (6 months–2 years) <i>Griffin 2018</i> ¹⁰³ <i>Mansell 2018</i> ¹⁰⁴ <i>Palmer 2019</i> ¹⁰⁵	6 people per 1000 had a SAE	29 more people had a SAE per 1000 (1 fewer to 283 more)	RR 6.21 (0.76 to 51.05)	348, 7 with SAEs	⊕⊕⊕⊕ Very low ⁵	RCT. Negative outcome, RR < 1 favours surgery. (<i>Griffin 2018</i>) An additional RCT reported no SAE in either treatment group (<i>Palmer 2019</i>) and <i>Mansell 2018</i> reported a single SAE which occurred in a patient who switched treatment arms

group to the arthroscopy group [as-treated analysis gave MD = -11.90 (-25.81 to 2.01); complete case: MD = 3.80 (-5.63 to 13.23)], which may help explain the heterogeneity. Very low certainty evidence.

Health-related quality of life at short-term follow-up (6 months–2 years) two randomised controlled trials^{103,104}

- Short-term (EQ-5D VAS): meta-analysis of 2 RCTs in 468 participants: MD (random effects) = 3.69 (-0.02 to 7.40), favouring surgery; EQ-5D VAS (0–100). Heterogeneity: I^2 = 0%. Low certainty evidence.

Sports activity at short-term follow-up (6 months–2 years) two randomised controlled trials^{104,105}

- Short-term: meta-analysis of 2 RCTs in 252 participants: MD (random effects) = -5.82 (-18.82 to 7.19); HOS sports subscale (0–100). Heterogeneity: I^2 = 68%. One study reported 70% of participants switched from the physiotherapy group to the arthroscopy group [as-treated analysis gave MD = -11.50 (-29.72 to 6.72); complete case: MD = 1.80 (-10.78 to 14.38)], which may help explain the heterogeneity. Low certainty evidence.

Adverse events and serious adverse events at short-term follow-up (6 months–2 years) three randomised controlled trials^{103–105}

- Three RCTs in 586 participants reported adverse and SAEs at short-term follow-up only.^{103–105}
- Griffin *et al.*¹⁰³ reported that among patients who received their allocated intervention, 100 of 138 patients in the osteochondroplasty group reported an AE versus 88 of 146 patients in the physiotherapy group: RR 1.20 (1.02–1.42); 284 participants, 188 with events. The most frequently reported AE was muscle soreness, reported by 58 (42%) patients in the hip osteochondroplasty group and 69 (47%) patients in the physiotherapy group. Six SAEs were reported in the surgical group and 1 in the physiotherapy group: RR 6.21 (0.76–51.05); 284 participants, 6 with events. Events in the surgical group were related to hospital re-admission, infection requiring antibiotic therapy or surgery and one falls incident. The SAE which occurred in the physiotherapy group was unrelated to treatment (biliary sepsis).
- Mansell reported 1 SAE (hip fracture) in 80 participants at short-term follow-up. The event occurred in a patient randomised to receive physiotherapy that switched to the surgical treatment arm.
- Palmer *et al.*¹⁰⁵ reported 3 AEs at short-term follow-up: RR 6.88 (0.36–131.58); 222 participants, 3 with events. All events occurred in the surgical group (superficial wound infection in one patient and injury to the lateral cutaneous nerve of the thigh in two patients). Authors reported that no SAEs occurred in either group.

Reoperation rate/revision of initial surgical procedure (including any new procedure) at short-term follow-up (6 months–2 years)

Two RCTs reported on this outcome at short-term follow-up only: RR 0.05 (0.01–0.23); 302 participants, 35 people with events. Very low certainty evidence. Most of these events we classified as 'new procedures' due to participants switching from their initial allocated treatment group of physiotherapy to the surgical treatment group.^{104,105}

- The Mansell 2018 trial reported 28 of the 40 patients (70%) assigned to physiotherapy eventually underwent surgery, 5 of these patients subsequently received revision surgery within the follow-up period. One THA was reported in the surgical treatment group.
- Palmer *et al.*¹⁰⁵ reported six people with events, all due to participants switching from conservative care to the surgical intervention group.

Progression at short-term follow-up (6 months–2 years)

- Short-term (2 years): RR 1.71 (0.44 to 6.67); 1 RCT in 79 participants, 1 with an event (development of hip OA at 2 years). NB: This study had 70% of participants switch from the physiotherapy to surgical intervention.¹⁰⁵
- Palmer *et al.*¹⁰⁵ report participants will be followed up for 3 years to evaluate the development of OA.

Type 2: comparison of non-regenerative surgical procedure versus another non-regenerative surgery

In the comparison of labral repair versus debridement, 1 RCT in 36 women reported medium-term outcomes (2.7 years' mean follow-up) for physical function and function during sports. Both interventions display an improvement from pre-operative baseline levels [e.g. debridement baseline 20.6 HOS ADL points (0–100 scale)]. Very low certainty evidence suggests there may be a benefit favouring labral repair compared to debridement, for example, a moderate-sized benefit of 10.3 HOS ADL (0–100) points (95% CI 0.37 to 20.2).

Considering osteochondroplasty versus lavage there is low certainty evidence from 1 RCT in 214 patients, suggests that both arthroscopic osteochondroplasty and lavage surgery (with or without labral repair for FAI) greatly improve combined pain and function from baseline by the 12 months' follow-up [(e.g. change from baseline for lavage is 37 iHOT-12 points (0–100)]. However, there may be no important difference between treatments for either pain and function, pain, function or QoL outcomes at short-term follow-up (12 months). For example, respective change scores for combined pain and function and pain at short-term; MD for iHOT-12 (0–100) = 2.20 (95% CI –6.95 to 11.35) and MD for VAS pain (0–100) = –4.40 (95% CI –13.41 to 4.61). At 2 years, very low certainty evidence suggests that the risk of reoperation may be lower in the osteochondroplasty group (RR 0.41, 95% CI 0.19 to 0.90), which corresponds to 110 fewer people having reoperations per 1000 (from 20 to 190 fewer) for a lavage group risk of 179 reoperations per 1000.

Type 3: comparison of non-regenerative surgical procedure versus another surgical procedure

There were no studies that met these criteria.

Conclusions

Femoro-acetabular impingement, a potential precursor to hip OA, is a target for non-regenerative procedures. Our review identified three RCTs and one NRS that compared non-regenerative surgical procedures with non-operative management, in patients with symptomatic FAI. The results suggest that this procedure does offer patients clinical relief of symptoms, above that seen with non-operative measures. The data of sufficient quality that enabled it to be taken forwards for health economic analysis.^{103–105,107}

The only other surgical procedure identified within the search was surgery to address damage to the hip labrum. However, there was no real evidence to support the use of labral reconstruction or labral repair, as a treatment for patients with early hip OA. This in part due to a lack of diagnostic clarity as in about one-quarter of the excluded studies the two surgical interventions had different indications for surgery. Therefore, we were encouraged to see the three new trials taking place to investigate the efficacy of labral surgery in patients with labral pathology and associated early knee OA.

In general, there was a lack of evidence of any other surgical treatments to treat early hip OA. Therefore, the focus of future research should be to gather further high-quality data on the two surgical procedures noted above.

Chapter 5 Load-modifying procedures to treat early osteoarthritis in the knee

Results of the search

The search and successive updates to January 2021 generated 6236 records. After deduplication and first assessment based on title and abstract screening, we obtained 117 studies in full text, comprising 30 systematic reviews of interventions and 74 potential RCTs or comparative NRS, and 13 potential prognostic factor studies in full text ([Figure 3](#)). The search strategy is recorded in [Appendix 1](#).

We included four intervention studies: three RCTs and one NRS.¹¹¹⁻¹¹⁴ There were no RCTs that directly compared a load-modifying procedure to non-operative management of early OA. There was one NRS that attempted to compare a load-modifying procedure to non-operative management of early OA. One RCT compared two load-modifying procedures. Two RCTs compared a load-modifying procedure to other types of surgery. We excluded 112 studies none of which could be used alone to answer the review questions.

There is one ongoing RCT study that does directly compare HTO with non-operative care.¹¹⁵ The protocol has been published but no results are available.

A further search of RCT studies published between January 2021 and January 2024 showed no other new RCTs of load-modifying procedures being compared to non-operative care.

Included studies characteristics

The main features of the characteristics of included studies are summarised in [Table 7](#). None of the 30 systematic reviews met our inclusion criteria sufficiently to be used as the sole evidence for this review without modification, so we conducted a de novo systematic review, drawing on the data in the systematic reviews. We obtained the full text of the included and excluded studies from the systematic reviews. We included three RCTs and one comparative NRS.¹¹¹⁻¹¹⁴

Characteristics of included studies

One NRS compared HTO with non-operative management through combining two other NRS.¹¹⁴ One study compared non-operative treatment versus an unloader brace.¹¹⁶ The second compared one type of HTO to another.¹¹¹ The second study is also reviewed alone as part of this review.

The other three studies compared two surgical procedures. One RCT compared two types of HTO load-modifying surgery.¹¹¹ One compared HTO with KJD surgery.¹¹³ The next compared load-modifying surgery with joint replacement surgery [HTO vs. unicompartmental knee arthroplasty (UKA)].¹¹²

Randomised controlled trial study sizes were generally small: 69, 72 and 92.¹¹¹⁻¹¹³ The NRS comprising 2 HTO comparisons had 113 and 99 participants in the 2 comparisons (drawn from a pool of 150 participants).¹¹⁴ No pre-published protocols were available for the RCTs.

The van Outeren study comprised two retrospective analyses of individual participant data (IPD) from two RCTs, comparing (a) HTO versus usual care (UC) and (b) UC versus unloader brace plus UC.¹¹⁴ The source RCTs compared opening versus closing wedge HTO and unloader brace plus UC versus UC. These trials were conducted during the same period in two different hospitals by the same research group and the first author of the trials is a coauthor of the van Outeren study.^{111,114,116} HTO participants from one trial were propensity score matched (one to many) with (a) UC participants or (b) brace plus UC participants from the other trial.

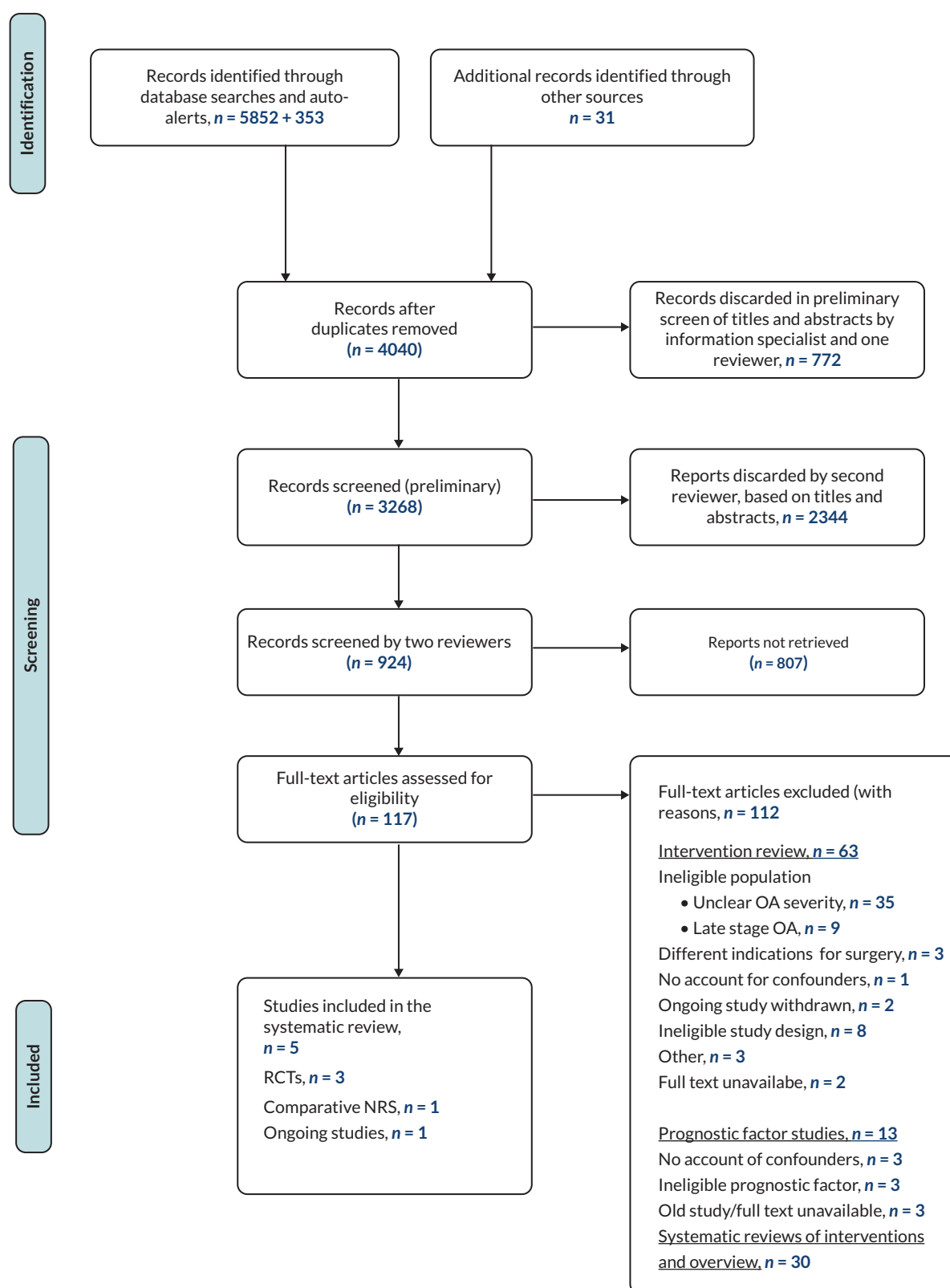


FIGURE 3 PRISMA diagram: load-modifying surgery for treatment of early knee OA.

TABLE 7 Summary of included load-modifying surgery studies characteristics for treatment of early knee OA

Study, design and country	Study size and duration	Population characteristics: mean age (SD), sex; male (%) BMI (kg/m ²)	Indication and stage of OA	Intervention	Comparator	Adjustment for confounding factors (method/factor adjusted for)	Outcomes
Type 1: load-modifying surgical procedure vs. non-operative management to directly treat patients with early knee OA (RCT and NRS)							
van Outeren 2017 ¹¹⁴ NRS – Prospective cohort, retrospective analysis of 2 RCTs The Netherlands	99 (71, 28) 12 months	<ul style="list-style-type: none"> Age: 51.7 (6.8) and 53.6 (11.2) Sex: 45 (63%) and 15 (54%) BMI: 28.6 (5.2) and 29.4 (5.5) 	<ul style="list-style-type: none"> HTO group: 1.1 (0.4) Ahlbäck Non-operative group: 1.2 (0.4) Ahlbäck 	HTO (mixed opening and closing wedges)	Non-operative management	<ul style="list-style-type: none"> Method: propensity score matching: one-to-many Factors: age, sex, hip-knee-angle, BMI, radiological severity of OA, baseline pain/HSS. Mixed-model analysis 	Pain and function, pain
Type 2: load-modifying surgical procedure vs. another load-modifying surgical procedure to directly treat patients with early knee OA (RCTs only)							
Brouwer 2006 ¹¹⁶ RCT (single-centre) The Netherlands	92 (47, 45) 12 months	Age: 50.2 Sex: 59 (64) and 33 (36)	OA Stage not detailed	HTO (closing wedge)	HTO (opening wedge)	RCT	Correction of deformity, pain, function
Type 3: load-modifying surgical procedure vs. another surgery to directly treat patients with early knee OA (RCTs only)							
Shaofei 2017 ¹¹² RCT (single-centre) China	72 (36, 36) 2 years	<ul style="list-style-type: none"> Age: 53.9 (7.4) and 52.5 (8.4) Sex: 3/36 (8%) and 5/36 (14%) BMI: NR Medial compartment	<ul style="list-style-type: none"> HTO group: KL2: 10/36 (28%) KL3: 26/36 (72%) UKA group: KL2: 11/36 (31%) KL3: 25/36 (69%) 	HTO	UKA	RCT	Combined pain and function, pain
van der Woude 2017 ¹¹³ RCT (multicentre) The Netherlands	69 (46, 23) 2 years	<ul style="list-style-type: none"> Age: 49.1 (1.0) and 51.2 (1.1) Sex: 27 (60%) and 16 (73%) BMI: 27.2 (0.5) and 27.5 (0.7) 	OA KL staging for both groups	HTO (opening wedge)	KJD	RCT	Pain and function, pain, function, QoL, progression
NR, not reported.							

Of the three RCTs, two were conducted in The Netherlands and one in China.¹¹¹⁻¹¹³ One RCT was conducted across two sites, and the other two were single-centre studies. No studies had industry funding and with all supported by public funding sources.

Population, intervention, control/comparison, outcome characteristics

These are recorded in [Table 7](#).

Surgical interventions assessed

The following surgical interventions were reported in the included studies:

- HTO (load-modifying surgical intervention): three RCTs and one NRS:
 - HTO (opening wedge): one RCT.^{112,113}
 - HTO (mixed opening and closing wedges): two NRS.¹¹⁴
- UKA: one RCT.¹¹²
- KJD: one RCT.¹¹³

Comparator interventions assessed

The following non-operative comparators were reported in the included studies:

- Non-operative management: standard care: that is patient education (adaptation of activities and/or weight loss), and (if needed) physical therapy and analgesics – this comparator represents NICE ‘core’ guidance.¹¹⁴
- Non-operative management: NICE ‘core’ guidance plus unloader brace.¹¹⁴

Outcomes

All studies reported at least one major outcome described in our protocol. Most outcomes related to pain or physical function, and combined pain and function, and used at least one validated instrument; details of outcomes, scales and time points are reported in [Appendix 1](#). Most studies only reported results at short-term follow-up (≤ 2 years). The following outcomes were reported:

- Pain: all RCTs.¹¹¹⁻¹¹⁴
- Physical function: one RCT.¹¹³
- Pain and function: two RCTs and one NRS.¹¹²⁻¹¹⁴
- PREMs: no studies reported this outcome.
- HRQoL: one RCT.¹¹³
- Subsequent Joint replacement (2 years’ follow-up or more): no studies reported this outcome.

Results

We present the results three sections. Type 1 studies: comparisons of load-modifying surgical procedures versus non-operative management; type 2 studies: comparisons of one load-modifying surgical procedure versus another load-modifying procedure and type 3 studies: comparisons of load-modifying surgical procedures versus other surgical options.

We report each outcome at short-, medium- and long-term follow-up for each of the individual comparisons, giving results separately for RCTs and the NRS. Risk-of-bias assessment for included studies is presented in [Appendix 2](#).

The summary of findings is presented in [Table 8](#). Only evidence from type 1 studies could be taken forward to the health economic analysis.

Type 1 studies: comparison of load-modifying surgical procedures versus non-operative management

There were no completed RCTs directly comparing load-modifying surgical procedures with non-operative treatment.

TABLE 8 Summary of findings table for load-modifying surgery studies for treatment of early knee OA

Load-modifying surgical procedure compared with non-operative management or another surgical procedure for early OA of the knee						
Patients or population: people with early OA of the knee				Settings: primary and secondary care		
Intervention: load-modifying surgical procedure				Comparison: non-operative management or another surgical procedure		
Comparisons of load-modifying surgical procedures vs. non-operative management						
Continuous outcomes	Score with non-operative management	MD	'Estimated scores' (95% CI)	Sample size (no. studies)	Evidence certainty (GRADE)	Comments
HTO vs. NICE core treatment (patient education – adaption of activities and/or weight loss, plus if needed physical therapy and analgesics) (non-randomised evidence)						
Pain at 12 months (short-term)	VAS pain score (0–10) = 5.2 (2.5) (baseline = 5.7)	MD: –1.60 (–2.71 to –0.49), that is moderate benefit favouring surgery	Authors' 'estimated' MD = –1.7 (–2.8 to –0.6), adjusted for baseline scores	99 (1 NRS)	⊕⊕⊕⊕ Very low ¹	NRS. Pain analysed as a negative outcome Negative SMD or MD favours surgery Propensity score matching of IPD from two RCTs Improvement in VAS pain (0–10) over baseline for non-operative group = 0.5
PROMs outcomes at medium-term and long-term duration – not measured						
Physical function outcome at short-term duration – not measured						
Pain and physical function at 12 months (short-term)	HSS (0–100) = 26.6 × (14.8) (baseline = 27.2)	MD: –6.20 (–12.49 to 0.09), that is small to moderate benefit-favouring surgery	Authors' 'estimated' MD = –6.6 (–13.1 to –0.2), adjusted for baseline scores	99 (1 NRS)	⊕⊕⊕⊕ Very low ¹	NRS. Pain analysed as a negative outcome Negative SMD or MD favours surgery Propensity score matching of IPD from two RCTs Improvement in HSS (0–100) over baseline for non-operative group = 0.6 (very small)
Adverse effects and reoperations not reported						

One NRS conducted an indirect analysis using data from 2 source RCTs to enable comparison of (a) HTO versus NICE core non-operative management – a combination of patient education (adaptation of activities and/or weight loss), and (if needed) physiotherapy and analgesics (99 participants), and (b) HTO versus unloader brace plus NICE core (113 participants) at 12 months' follow-up.¹¹⁴ One RCT compared two types of HTO, and the other compared brace plus NICE core versus NICE core. The only outcomes reported were pain and combined pain and function. For HTO versus NICE core treatment, very low certainty evidence suggests there may be a moderate benefit in each of pain and combined pain and function, favouring the HTO group: for example, the MD for VAS pain (0–10) was –1.6 (95% CI –2.7 to –0.5). The NICE core group had only a small improvement over baseline in PROMs values at 12 months' follow-up (e.g. 0.5 VAS pain points). For the comparison of HTO with brace plus NICE core, there was a moderate benefit for HTO compared with brace [MD = –1.2 (–2.3 to –0.08)] on the VAS pain scale of 0–10. However, this time there was also a moderate improvement in the brace group for pain compared to values preoperatively (1.4 VAS pain points). There was little difference between interventions for the pain and function outcome, although both groups improved over baseline [7.4 HSS points (0–100) for the brace group].

This was a composite indirect study, so the results are not presented in a more detailed manner. The evidence from this study was too weak to be taken forward to the health economic analysis due to its indirect nature.

Type 2 studies: comparison of one load-modifying surgical procedure versus another load-modifying procedure or other joint-preserving surgical procedures

There was one RCT that compared opening versus closing wedge HTO.¹¹¹ The primary outcome measure was correction of alignment. Functional outcome at 1 year post operation was a secondary outcome. There was no difference between groups at 1 year on the VAS pain scale of 0–10 and the HSS knee score. However, no preoperative data were provided so the improvement from baseline is not available.

Type 3 studies: comparison of load-modifying surgical procedure versus joint replacement surgery

The Shaofei RCT compared HTO versus joint replacement surgery (UKA).¹¹² Again, the evidence is of very low certainty. RCT evidence from 1 study in 72 participants suggests there may be a moderate benefit for UKA compared to HTO at 2 years' follow-up for pain, and for combined pain and function. For example, for pain and function (HSS knee score, 0–100), the MD was 10.3 (95% CI 6.4 to 14.2), favouring UKA, and the change from baseline for HTO was 28 HSS points. At long-term follow-up, the NRS findings were inconclusive for PROMs outcomes and both NRS had features that made the findings unreliable. Very low certainty evidence for a meta-analysis of the results from these 2 studies in 350 participants for the reoperation outcome suggests the hazard for HTO may be about twice that for UKA for conversion to TKR at long-term follow-up (up to 30 years) – hazard ratio (HR) = 2.13 (95% CI 0.79 to 5.56), with some heterogeneity in the results. This corresponds to 115 more people per 1000 who had received HTO (rather than UKA) having revisions to TKR, from 53 fewer to 185 more.

One small RCT in 69 participants compared HTO with KJD at short-term follow-up (12 months and 2 years).¹¹³ Evidence is of very low certainty, mainly because of high risk of bias and imprecision. Both surgical interventions substantially improved pain and physical function compared with their own preoperative values (e.g. 38 points for HTO on VAS pain (0–100) scale). For HTO versus KJD, there was a small decrease in pain of nine VAS points on a 0–100 scale, favouring HTO at 2 years (95% CI –3.8 to 21.8), but the CI is wide and a large difference in baseline between the two interventions made the results less reliable. The evidence suggests there may be little difference in pain and function at 2 years' follow-up (MD = –1.4, 95% CI –10.8 to 8.0), but again there is a large difference in baseline values compared with the effect estimate. The RCT reported that 59% of KJD participants had single or multiple pin tract infections and 4% of HTO participants had postoperative wound infections.

Conclusions

This review of load-modifying procedures in people with early knee OA identified a very limited evidence base of comparative studies of surgical interventions versus non-operative treatment.

The load-modifying procedure with the most published evidence is HTO. However, most of the evidence we identified for HTO load-modifying surgery in early OA is informed by single studies, and is of very low certainty, mainly because of risk of bias and imprecision in the estimates. Only one NRS informed the question on load-modifying surgical procedures versus non-operative management, but this was an indirect study. The study compared HTO with an unloader brace and with UC, in a study that combined IPD from two separate source RCTs. As such the indirect evidence provided is of low quality, as no direct comparison between surgery and non-operative treatment is presented.

The other studies, all RCTs, include comparisons of HTO to other surgical procedures. One compares different ways of performing a HTO, the second compares HTO to UKA and the third studies HTO compared to KJD. They provide no direct evidence of the efficacy of any load-modifying surgery when compared to non-operative measures.

One ongoing study was identified that does directly compare HTO to non-operative treatment and the results of this will be an important contribution to the literature.

Knee joint distraction is a procedure that has not been widely studied. There is only one published study in this report, which is a RCT comparing KJD with HTO.¹¹³ Our search revealed a further RCT study of KJD compared to TKR. Future study will be required to identify if this treatment is effective in treating early knee OA.

In summary, there were very little data that offered direct comparison between HTO and non-operative measures. As a result, there were no data of sufficient quality to take forward to a health economic analysis.

Chapter 6 Load-modifying surgical procedures around used to treat early hip osteoarthritis

Results of the search

The search and successive updates to January 2021 generated 6474 records. After deduplication and first assessment based on title and abstract screening, we obtained 56 reports in full text, comprising 12 systematic reviews, 36 reports of potential comparative NRS and 5 prognostic factor studies, 1 a meta-regression ([Figure 4](#)). The search strategy is recorded in [Appendix 1](#).

There were no RCTs or NRS comparing PAO against non-operative treatment. There were no randomised controlled studies comparing different types of PAO. There were no RCTs that compared PAO to another surgical treatment.

In view of this lack of evidence, we included a number of non-randomised comparative studies that related to PAO.

There were two NRS that compared PAO against another surgical treatment, which would normally have been excluded. One study was presented in as a conference abstract in the grey literature, which compared PAO against PAO plus hip arthroscopy with hip arthroscopy alone.¹¹⁷ The second study compared PAO to THR.¹¹⁸

In addition, we included 10 comparative studies examining the outcome for THR performed in patients who had previously undergone PAO or patients who had not previously undergone PAO.¹¹⁹⁻¹²⁸

In addition, we assessed two systematic reviews of non-comparative PAO cohort studies who would have been excluded from the search but we felt were of interest, given the overall lack of evidence.^{129,130} These studies were not eligible to have data examined in detail or to provide data for the health economic analysis.

We are aware of one ongoing study (NCT 03941171). This is a RCT that compares osteotomy with non-operative management in people with hip dysplasia scheduled for PAO. The study aims to examine the efficacy of PAO followed by 4 months of UC followed by 8 months of progressive resistance training (PRT) compared with PRT alone for 12 months. The study is due to report results at one, 5 years and 10 years follow-up and the primary completion date is January 2025.¹³¹

Included studies characteristics

The main features of the characteristics of included studies are summarised in [Table 9](#).

Study details

Within the search criteria we found: no RCTs or NRS that compared load-modifying surgery versus non-operative management, no RCTs that compared versus another type of load-modifying surgery and no RCT studies that compared load-modifying surgery versus non-operative versus another type of surgery.

One ongoing RCT was identified that compared load-modifying surgery versus non-operative management.¹³¹

In the absence of comparative studies for load-modifying surgery, we included a number of other studies that we believed would be of interest, acknowledging that they could not inform the health economic analysis.

We included two NRS that compared PAO to other surgical interventions.^{117,118} We included 10 cohort studies that assessed the outcome of THR after PAO.¹¹⁹⁻¹²⁸

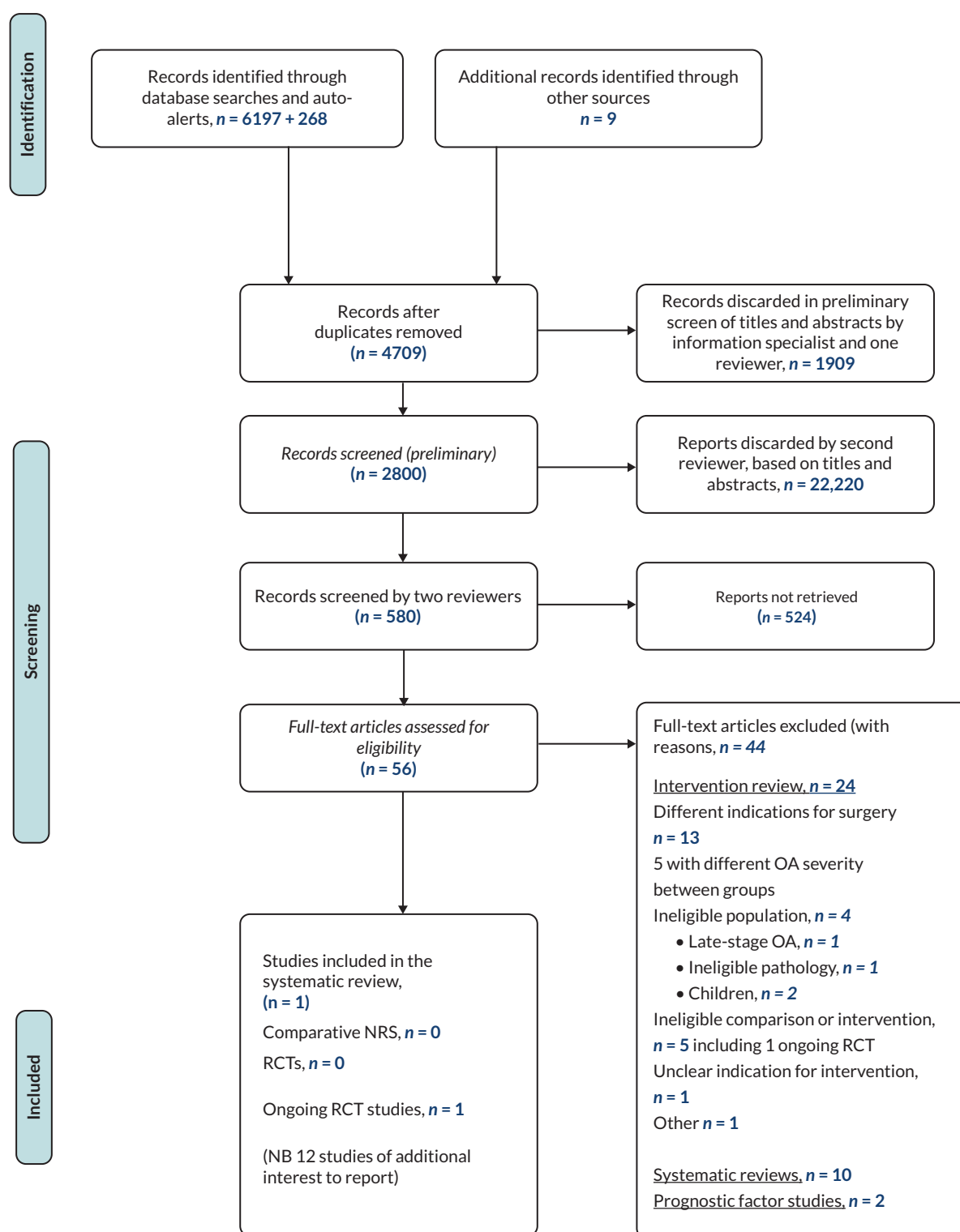


FIGURE 4 PRISMA diagram: load-modifying surgery for treatment of early hip OA.

In addition, assessed two recent systematic reviews of non-comparative studies to give information on (1) change from baseline PROM scores, and (2) the number of participants who went on to have a THR, in people who had osteotomy for dysplasia.^{129,130}

Characteristics of included studies

Two comparative studies of PAO against another surgical intervention were included. Kenney *et al.* compared PAO with arthroscopy against arthroscopy alone in a study of 124 patients with hip dysplasia in the USA.¹¹⁷ Gray *et al.* compared

TABLE 9 Summary of included load-modifying surgery studies characteristics for treatment of early hip OA

Study ID, design and country	Study size and duration	Population: mean age (SD), sex; male (%), BMI (kg/m ²)	Pathology and indication for surgery (KL/Tönnis grade)	Intervention	Comparator	Adjustment for confounding factors (method/factors adjusted for)	Outcomes
<i>Type 1: comparison of load-modifying surgical procedures vs. non-operative management (RCTs and NRS)</i>							
No studies identified							
<i>Type 2: comparison of load-modifying surgical procedures vs. another load-modifying procedure (RCTs)</i>							
No studies identified							
<i>Type 3: comparison of load-modifying surgical procedures vs. another surgical procedure (RCTs)</i>							
No studies identified							

the results of PAO against THR in 155 patients in the USA. The indication for each surgery was different with patients with hip dysplasia receiving PAO and patients with OA receiving arthroplasty. This paper was reported as a conference proceedings abstract, and the study has not been published elsewhere.

A total of 10 retrospective NRS reported the outcome of hip arthroplasty in patients who had previously undergone PAO. For these studies, the median sample size was 87 (IQR 64.5–120.5); 4 studies included 100 or more participants.^{122,123,125,127}

Of the 10 NRS, 4 studies were conducted in Japan, 3 in the USA and single NRS was completed in Canada, France and UK.^{119–128}

Two received no funding, one study had commercial funding, two had a mixture of commercial and non-commercial funding. For the other studies, funding was either unclear or did not report any funding source.

Population, intervention, control/comparison, outcome characteristics

There were no relevant PICO to assess in these studies.

Surgical interventions assessed

The following surgical interventions, stratified as above, were reported in the included studies.

- PAO (type not specified): one NRS.¹¹⁸
- PAO (type not specified) plus arthroscopy (no details): one NRS.¹¹⁷
- History of osteotomy: 10 NRS:
 - Rotational acetabular osteotomy (RAO).^{120,122,128}
 - One each of Bernese.^{119,121,126}
 - Mixed PAO types: three NRS.^{124,125,127}
 - Both pelvic and femoral osteotomy types.¹²³

Comparator interventions assessed

- No history of osteotomy: 10 NRS.^{119–128}
- Arthroscopy (unspecified): one NRS.¹¹⁷
- THA: one NRS.¹¹⁸

No studies explicitly reported NICE 'core' guidance as a comparator (combination of information, exercise/physical therapy and weight loss) or the individual components.

Outcomes

All studies reported at least one major outcome described in our protocol. Most outcomes related to combined pain and function, AEs and reoperation/revision rates and used at least one validated instrument.

One study comparing osteotomy plus arthroscopy versus arthroscopy reported results at short-term follow-up (6 months–2 years).¹¹⁷ The other studies investigated the impact of prior osteotomy (vs. none) in people who subsequently had a THR, and reported results either at medium-term follow-up post THR (> 2 years to < 5 years) or at long-term follow-up post THR (≥ 5 years). The following outcomes were reported by the studies:

- Pain and function: 12 NRS.^{118–129}
- Pain: six NRS.^{117,120,122–125}
- Physical function: six NRS.^{117,120,122–125}
- HRQoL: three NRS.^{124,125,127}
- Surgical complications and AEs: 10 NRS.^{119–128}
- Revision of THA (2 years' follow-up or more): 10 NRS.^{119–128}
- Reoperation rate/revision of initial surgical procedure, including any new procedure: five NRS.^{119,121,122,125,127}

Results

We present the results of three types of procedures: type 1 studies: comparisons of load-modifying surgical procedures versus non-operative management; type 2: comparison of load-modifying surgical procedures versus another load-modifying procedure and type 3: comparisons of load-modifying surgical procedures versus other surgical procedures.

Type 1: comparison of load-modifying surgical procedures versus non-operative management

No RCTs or comparative NRS was identified that directly compared a load-modifying procedure to non-operative treatment.

Type 2: comparison of load-modifying surgical procedures versus another load-modifying procedure

No RCTs were identified that directly compared a load-modifying procedure to non-operative treatment.

Type 3: comparison of load-modifying surgical procedures versus another surgical procedure

No RCTs were identified that directly compared a load-modifying procedure to another surgical procedure.

Additional analysis

As there were no studies identified above, we included 12 additional studies outside the scope of the search parameters that we felt may be of interest ([Table 10](#)).

There were two studies comparing PAO to other types of surgery that offer some direct information regarding the outcome of PAO. We identified in the grey literature 1 NRS in 124 participants comparing osteotomy plus hip arthroscopy versus hip arthroscopy alone (no details were given about the arthroscopy).¹¹⁷ This study is published as two conference abstracts and presentation slides, and reports short-term (2-year) outcomes of pain, physical function and combined pain and function. We assessed this study to be at high risk of bias for many domains, and it was unclear whether the authors had attempted to take any account of confounding factors in the analysis or study design. Possibly, because of this, two of the three outcomes had very large differences between interventions preoperatively, such that the effect estimate was dominated by the baseline differences, making the findings potentially unreliable. We report here the remaining outcome, pain – sounding a note of caution about these results too. Very low certainty evidence suggests a moderate to large benefit favouring osteotomy plus arthroscopy over arthroscopy of 1.60 VAS pain points on a scale of 0–10 (95% CI 0.55 to 2.65). Both treatments led to an improvement in pain compared with baseline values (arthroscopy giving an improvement of 2.8 points). The second study compares the outcome of patients who underwent PAO for hip dysplasia, with a group of patients who underwent THR for hip OA.¹¹⁸ The study reports that patients' hip

TABLE 10 Summary of additionally included studies of load-modifying surgery for treatment of early hip OA

Study ID, design and country	Study size and duration	Population: mean age (SD), sex; male/female (%), BMI (kg/m ²)	Pathology and indication for surgery (KL/Tönnis grade)	Intervention	Comparator	Adjustment for confounding factors (method/factors adjusted for)	Outcomes
<i>Comparison of load-modifying surgical procedures vs. another surgical procedure (NRS and Grey literature)</i>							
Kenney 2018 ¹¹⁷ (Abstract) Non-randomised comparative study of 2 cohorts USA	124 (34,90) 1.8 and 3.4 years	<ul style="list-style-type: none"> Age: 31 (range 15–51) and 32 (range 14–62) years Sex: 5 (15%) and 32 (36%) BMI: NR 	Dysplasia No details	PAO + arthroscopy (no details)	Arthroscopy (no details)	<ul style="list-style-type: none"> Method: unclear Factors: apparently not matched. Baseline difference in sex (15% and 35% male) and large differences in preoperative PROMs scores for HHS and HOS ADL. Age similar 	Pain, function, Pain and function
Gray 2015 ¹¹⁸ NRS– comparison of two cohorts USA	155 (100, 55) 5.9 years (2–13)	<ul style="list-style-type: none"> Age: 28.9 (18–40) and 32.7 (18–40) Sex: 26% and 36% BMI: 25.4 (18–34) and 27.2 (18–35) 	Dysplasia (PAO group) OA (THA group)	PAO	THA	<ul style="list-style-type: none"> Retrospective No adjustments 	Pain, function, activity
<i>Comparison of THR outcome in patients who have previously undergone PAO or those who have not previously undergone PAO</i>							
Amanatullah 2015 ¹¹⁹ Retrospective cohort study Multicentre (2 centres) USA	53 (30, 23) hips 10 years (4) and 6 years (4). Minimum 2 years	<ul style="list-style-type: none"> Age: 38 (11) and 38 (10) years Sex: 6 (26%) and 6 (26%) BMI: 29 (6) and 28 (7) 	Dysplasia <ul style="list-style-type: none"> OA severity: NR Pre-PAO: Tönnis angle= 26° (8); anterior CEA= 1° (15); LCEA= 4° (9); all participants Crowe grade 1 Post PAO: Tönnis angle= 15° (11); anterior CEA= 28° (23); LCEA= 20° (15); 22 patients (96%). Crowe grade I, 1 patient Crowe grade II (4%) 	History of PAO (Bernese osteotomy)	No prior surgery	<ul style="list-style-type: none"> Method: simple matching Factors: diagnosis (DDH), age, sex, BMI 52 control group patients available for matching: took first available match, without a history of PAO and a minimum 2 years of follow-up after THA to give 1 : 1 matching 	Pain and function, AEs, revision/reoperation
Fukui 2015 ¹²⁰ Retrospective cohort study Single-centre (assumed) Japan	52 (22, 30) hips 8.2 years (7–11 range) and 8.7 years (7–11)	<ul style="list-style-type: none"> Age: 53.6 (46–62 range) and 55.8 (46–71) years Sex: 2 (10%) and 2 (7%) BMI: 23.1 (SD NR) and 23 (SD NR) 	Dysplasia <ul style="list-style-type: none"> OA severity: NR Group 1 – Crowe grade 1: 18 (82%); grade 2: 4 (18%). LCEA before THA: 27 (3–42 range) Group 2 – Crowe grade 1: 27 (90%); grade 2: 3 (10%). LCEA: 16 (10–23 range) 	History of PAO (RAO)	No prior surgery	<ul style="list-style-type: none"> Method: simple matching Factors: stated to be ‘diagnosis-matched’ (DDH assumed); similar at baseline for age, sex, BMI, Crowe and HHS, but different for LCEA 	Pain and function, radiological outcomes, AEs (pain, function measured but no results reported)

Study ID, design and country	Study size and duration	Population: mean age (SD), sex; male/female (%), BMI (kg/m ²)	Pathology and indication for surgery (KL/Tönnis grade)	Intervention	Comparator	Adjustment for confounding factors (method/factors adjusted for)	Outcomes
Hashemi-Nejad 2002 ¹²¹ Retrospective cohort study Single-centre UK	83 (33, 50) hips 60 (range 24–199) months; that is 5 years (2–16.6)	<ul style="list-style-type: none"> Age: 45 (38–64 range) and 39 (25–51) years Sex: 2 (8%) and 4 (9%) BMI: NR 	Dysplasia. <ul style="list-style-type: none"> OA severity: NR for sample, but original study reported mean pre-osteotomy score of 0.7 (scale NR, but 0 = no OA and 1 is marginal osteophytes); after 10 + years, mean score was 2.7 (2 = minimal loss of joint space, 3 = obvious JSN and sclerosis) Group 1 – Crowe grade 1: 20 (71%); grade 2: 7 (25%); grade 3: 1 (4%) Group 2 – Crowe grade 1: 36 (72%); Crowe grade 2: 8 (16%); Crowe grade 3: 4 (8%); Crowe grade 4: 2 (4%) 	History of PAO (Chiari)	No prior surgery	<ul style="list-style-type: none"> Method: simple matching Factors: said to be 'diagnosis-matched'; similar at baseline for sex, side of THA, Crowe classification, follow-up time; significantly different for age (6 years younger in control group) 	Pain and function, reoperation/revision, AEs, PREMs
Ito 2015 ¹²² Retrospective cohort study Single-centre Japan	102 (44, 58) hips Mean 55.8 and 62.9 months	<ul style="list-style-type: none"> Age at THA: 55.6 (7.8) years Age at RAO: 34 (12.4) and 56.2 (5.1) years Sex: 2 (5%) and 2 (3%) BMI: 22.8 (3.4) and 22.3 (2.7) 	Dysplasia <ul style="list-style-type: none"> OA severity: NR Group 1 – Crowe grade 1: 29 (66.0%); grade 2: 10 (22.7%); grade 3: 3 (6.8%); grade 4: 2 (4.5%) Group 2 – Crowe grade 1: 39 (67%); grade 2: 14 (24%); grade 3: 4 (7%); grade 4: 1 (2%) 	History of PAO (RAO)	No prior surgery	<ul style="list-style-type: none"> Method: simple matching Factors: matched for age and sex (similar baseline for BMI, preoperative PROMs and Crowe classification) 	Pain and function, pain, function (walking), revision/reoperation, AEs
continued							

TABLE 10 Summary of additionally included studies of load-modifying surgery for treatment of early hip OA (*continued*)

Study ID, design and country	Study size and duration	Population: mean age (SD), sex; male/female (%), BMI (kg/m ²)	Pathology and indication for surgery (KL/Tönnis grade)	Intervention	Comparator	Adjustment for confounding factors (method/factors adjusted for)	Outcomes
Migaud 2014 ¹²³ Retrospective cohort study Multicentre France	430 (159, 271) hips 13.2 (5.4) years for PROMs and 15-year survivorship analysis	<ul style="list-style-type: none"> Age: 55 (12) and 57 (10) years Sex: 19 (15%) and 44 (22%) BMI: 23.9 (3.8) and 24.5 (4.4) 	Dysplasia <ul style="list-style-type: none"> OA severity: NR Group 1 – Crowe grade 1: 29 (20%); grade 2: 35 (24%); grade 3: 32 (22%); grade 4: 49 (34%) and <i>n</i> = 14 missing data Group 2 – Crowe grade 1: 56 (23%); grade 2: 62 (26%); grade 3: 40 (16%); grade 4: 83 (35%) and <i>n</i> = 30 missing data 	History of osteotomy (mixed osteotomy: PAO and femoral osteotomy)	No prior surgery	<ul style="list-style-type: none"> Method: other Factors: said to be 'comparable' at baseline (pre-THA) for age, sex, Devane Activity Score, Charnley Hip Score, Crowe classification. Revision outcome may be analysed taking into account of covariates (mentions first-order risk threshold of 5%) 	Pain and function, pain, function, revision/reoperation, survivorship, AEs
Osawa 2018 ¹²⁴ Retrospective cohort study Multicentre (partly): also included 6/29 intervention group hips from other hospitals. Japan	88 [30 (1 of whom died), 58] hips 3.0 and 3.1 years	<ul style="list-style-type: none"> Age: 57.2 (7.2) and 58.3 (7.7) years Sex: 2 (7%) and 4 (7%) BMI: 24.2 (3.6) and 23.6 (3.3) 	Dysplasia <ul style="list-style-type: none"> OA severity: NR Group 1 – Crowe grade 1: 16 (55%); grade 2: 11 (38%); grade 3: 2 (7%); grade 4: 0 Group 2 – Crowe grade 1: 32 (55%); grade 2: 22 (38%); grade 3: 4 (7%); grade 4: 0 	History of PAO [PAO (mixed)]	No prior surgery	<ul style="list-style-type: none"> Method: simple matching Factors: matched 1 : 2 on age ($\pm$ 5 years), sex, and Crowe classification (and similar at baseline for BMI and preoperative scores) 	Pain, pain and function, QoL, AEs, PREMs
Stambough 2015 ¹²⁵ Retrospective cohort study Multicentre (7 of 11 possible centres) USA	117 (39, 78) 2.2 and 2.5 years	<ul style="list-style-type: none"> Age: 19.5 (7), range 10.9–35.4 and 19.9 (6), range 9–35.4 years Sex: 4 (10%) and 8 (10%) BMI: 23.3 (4.8) and 24.0 (4.3) 	Dysplasia <ul style="list-style-type: none"> Group 1 – OA severity: Tönnis grade 0: 15 (38%); grade 1: 23 (59%); grade 2: 1 (3%). Crowe grade: NR; LCEA mean: 10 (range –29–50) Group 2 – Tönnis grade 0: 38 (49%); grade 1: 38 (49%); grade 2: 2 (3%). Crowe grade: NR, LCEA mean 7 (range –30–50) 	History of osteotomy [PAO (mixed)]	No prior surgery	<ul style="list-style-type: none"> Method: simple matching Factors: age (within 1 year), sex, and BMI (within 4 kg/m²). 1 : 2 ratio of intervention: control. Some differences in Tönnis score (study group: 38% grade 0, 59% grade 1; control: 49% grade 0, 49% grade 1). Matching by an author blinded to outcomes. 	Pain, function, pain and function, QoL, AEs, THA conversion

Study ID, design and country	Study size and duration	Population: mean age (SD), sex; male/female (%), BMI (kg/m ²)	Pathology and indication for surgery (KL/Tönnis grade)	Intervention	Comparator	Adjustment for confounding factors (method/factors adjusted for)	Outcomes
Tokunaga 2011 ¹²⁶ Retrospective cohort study Single-centre (assumed, all surgery performed by senior author) Canada	103 (52, 51) hips 7.3 and 8.5 years	<ul style="list-style-type: none"> Age: 41 (25–58 range) and 47 (27–66 range) years Sex: 2 (4%) and 2 (5%) BMI: NR 	Dysplasia <ul style="list-style-type: none"> OA severity: NR. Group 1 – Crowe grade 1: 48 (92.3%); grade 2: 4 (7.7%) Group 2 – Crowe grade 1: 46 (90.2%); Grade 2: 5 (9.8%) 	History of PAO (Salter and Chiari)	No prior surgery	<ul style="list-style-type: none"> Method: baseline similarities Factors: data set split according to intervention – no planned account of confounding factors. Results showed similarities at baseline (pre-THA) for sex, preoperative PROMs scores, Crowe classification. BMI NR. Age significantly lower by 6 years for PAO group. 	Pain and function, revision/reoperation, survivorship up to 10 years (THA)
Yacovelli 2020 ¹²⁷ Retrospective cohort study Multicentre (2 centres) for osteotomy but all THA conducted at the same institution USA	196 (49, 147) 3.8 (2.6) and 3.8 (2.1) years (minimum 2 years)	<ul style="list-style-type: none"> Age: 38.0 (11.2) and 41.7 (10.0) years Sex: 6 (12.2%) and 22 (15.0%) BMI: 27.4 (6.2) and 26.8 (6.2) 	Dysplasia <ul style="list-style-type: none"> OA severity: NR. Both groups – Crowe type 1 and 2 but individual breakdown is NR 	History of PAO [PAO (mixed)]	No prior surgery	<ul style="list-style-type: none"> Method: propensity score matching: 1 : 3 Factors: propensity score matched on age, sex, BMI and date of surgery (however, mean age 3.7 years difference between groups). NB difference in HOOS preoperative score of 6.4 points (effect estimate 7.0) and 3.1 SF-36 points (effect estimate 0.7) 	Pain and function, revision/reoperation, progression, AEs
Yuasa 2015 ¹²⁸ Retrospective cohort study Single-centre Japan	48 (24, 24) hips 85 (15–195 range) months and 86 (16–192 range) months	<ul style="list-style-type: none"> Age: 57.1 (9.8) and years 59.8 (8.2) Sex: 3 (14%) and 4 (17%) BMI: 23.4 (1.85) and 23.0 (2.94) 	Dysplasia <ul style="list-style-type: none"> OA severity: NR. 	History of PAO (RAO)	No prior surgery	<ul style="list-style-type: none"> Method: simple matching Factors: age and diagnosis (sex and BMI, and pre-operative HHS also similar at baseline) 	Pain and function, revision/reoperation, AEs
NR, not reported.							

pain and function both improved after both procedures. The PAO patients improved from a preoperative mean modified Harris Hip Score (mHHS) of 62.5 (SD 2.5) to 86.5 (SD 14.8) postoperatively. However, the mean change score was less than the THR patients, with 31.1 (SD 20.8) compared to 24 (SD 16.3). Neither study offered any comparative evidence of the outcome of PAO compared to non-operative treatment.

Further indirect evidence came from studies comparing the outcome of THR in patients who had previously undergone a PAO versus with those who had not. The evidence was from the 10 NRS and was of very low certainty.¹¹⁸ Most studies attempted to allow for key confounding factors using simple matching of baseline characteristics, and one used propensity score matching. As expected, there were large differences between PROMs scores pre- and postoperatively to THR, regardless of whether participants had a history of osteotomy. For the outcome of combined pain and function at medium-term follow-up (more than 2 years to < 5 years), very low certainty evidence from meta-analysis of results from three NRS in 385 hips suggests a small *worsening* in outcome for those with a history of osteotomy compared to those without osteotomy: SMD 0.44 (95% CI 0.10 to 0.77), with some heterogeneity. This SMD was back-transformed to give a mean difference on the HHS scale: MD was 3.36 (95% CI 0.76 to 5.88) on a scale of 0–100. Findings are similar at long-term follow-up (5 years or more): very low certainty evidence from 5 NRS in 484 hips suggests a slightly poorer outcome in those with a history of osteotomy: meta-analysis gave a very small worsening in the HHS of 2.12 (95% CI 0.31 to 3.94) points, with no heterogeneity among studies.

Given the lack of outcome data regarding PAO, we also conducted additional analyses of the data reported in two recent systematic reviews of before-and-after studies in people who had PAO surgery. These studies were excluded from our search at the full-text assessment stage, as they do not include any studies where comparison of PAO to non-operative treatment was made.^{129,130} We have not quality assessed the reviews or the component studies. We stress that these results only concern the change post- versus preoperative to osteotomy and tell us nothing about the comparison with non-operative management or other surgeries. However, these analyses do give some evidence regarding the potential benefits and harms following osteotomy and indicate the effect of OA severity on the success of osteotomy. Our results suggest that, first, osteotomy may give moderate to large improvements in pain and function over baseline of between 11 and 29 HHS points (on a scale of 0–100) at 2 years' follow-up or more, with improvements in pain and function correlating inversely with duration of follow-up (data from 26 cohorts). Meta-analysis gave a MD (random effects) change from baseline score of 20.7 HHS points (95% CI 18.8 to 22.6), with a high degree of heterogeneity ($I^2 = 88\%$), but results were in the same direction. The second study examined 44 studies including 4790 hips, of which 412 were converted to THR, suggesting that the median (IQR) proportion of osteotomies converted to THA is 6.5% (2.4–11.5), with mean 8.6% at a follow-up of 2 years and above. This conversion rate increased with duration of follow-up, with mean values of 4.5% at 2 years to < 5 years (21 studies), 10.1% at 5 years to < 10 years (11 studies) and 15.7% at 10 years or more follow-up (10 studies).

Conclusions

Our search revealed that PAO is the only established treatment that is used in this field of care, but there is deficit of randomised controlled evidence to support its use. There is some weak evidence of its clinical efficacy, but the real treatment effect of the procedure, when compared to non-operative treatment, has not been established.

The additional analysis we performed did identify cohort studies that suggest that PAO may offer patients improvements in pain and function. This work suggests further research is needed to understand its true effect and its comparison to non-operative treatment.

We are aware of one ongoing RCT that compares the clinical effect of PAO compared to non-operative treatment in patients with hip dysplasia undergoing PAO. The trial is under way but has not reported yet, but appears to be methodologically sound. The results of this study will offer much needed evidence to understand the outcome of PAO. In addition, we would recommend that there is investment in further RCTs to investigate this mode of treatment.

In summary, this study identified a lack of any high-level evidence of the efficacy of load-modifying procedures to treat early hip OA. Therefore, no data could be taken forward to the health economic analysis.

Chapter 7 Biological surgical reconstruction for early knee osteoarthritis

Results of the search

The search and successive updates up to January 2021 generated 7979 records ([Figure 5](#)). After deduplication, this number reduced to 5189. Further screening was performed initially by title and abstract, then by full-text assessment. The search strategy is recorded in [Appendix 1](#). We found no studies that met our search criteria that compared a surgical technique against non-operative treatment. We included only one study, which met the criteria of a comparative study of patients being treated for early OA, which compared one biological reconstruction technique to another.¹³²

We did identify a series of RCTs that studied a number of different surgical techniques for biological reconstruction of articular cartilage defects.¹³³⁻¹⁵⁴ Of the 23 studies one was in progress with no published results.¹⁵⁵ None of these studies could be defined as a direct or indirect treatment of OA. They are included as they were considered of interest to this piece of work as each method is a potential treatment for OA.

Included studies characteristics

The main features of the characteristics of included studies are summarised in [Tables 11](#) and [12](#).

Study details

Only one study could be included in the systematic review.¹³² This RCT is the only study identified that studies biological reconstruction in patients with early OA, as defined in our study.

Characteristics of included studies

The study does not compare a biological repair technique with non-operative treatment. It compares AMIC with MF. The study was completed in South Korea and was funded by industry.

Population, intervention, control/comparison, outcome

The PICO data for the included trial are presented in [Table 11](#).

Surgical interventions assessed

The following surgical interventions were reported in the included studies:

- MF
- AMIC

Comparator interventions assessed

No studies compared a surgical intervention to any form of non-operative care package for OA. In the single study included MF was the comparator treatment.

Outcomes

The Kim study used the KOOS score and a VAS score, with both outcome measures described in our protocol.¹³²

Results

We present the results of three types of procedures: type 1 studies: comparisons of biological surgical reconstruction procedures versus non-operative management; type 2: comparisons of biological surgical reconstruction versus another

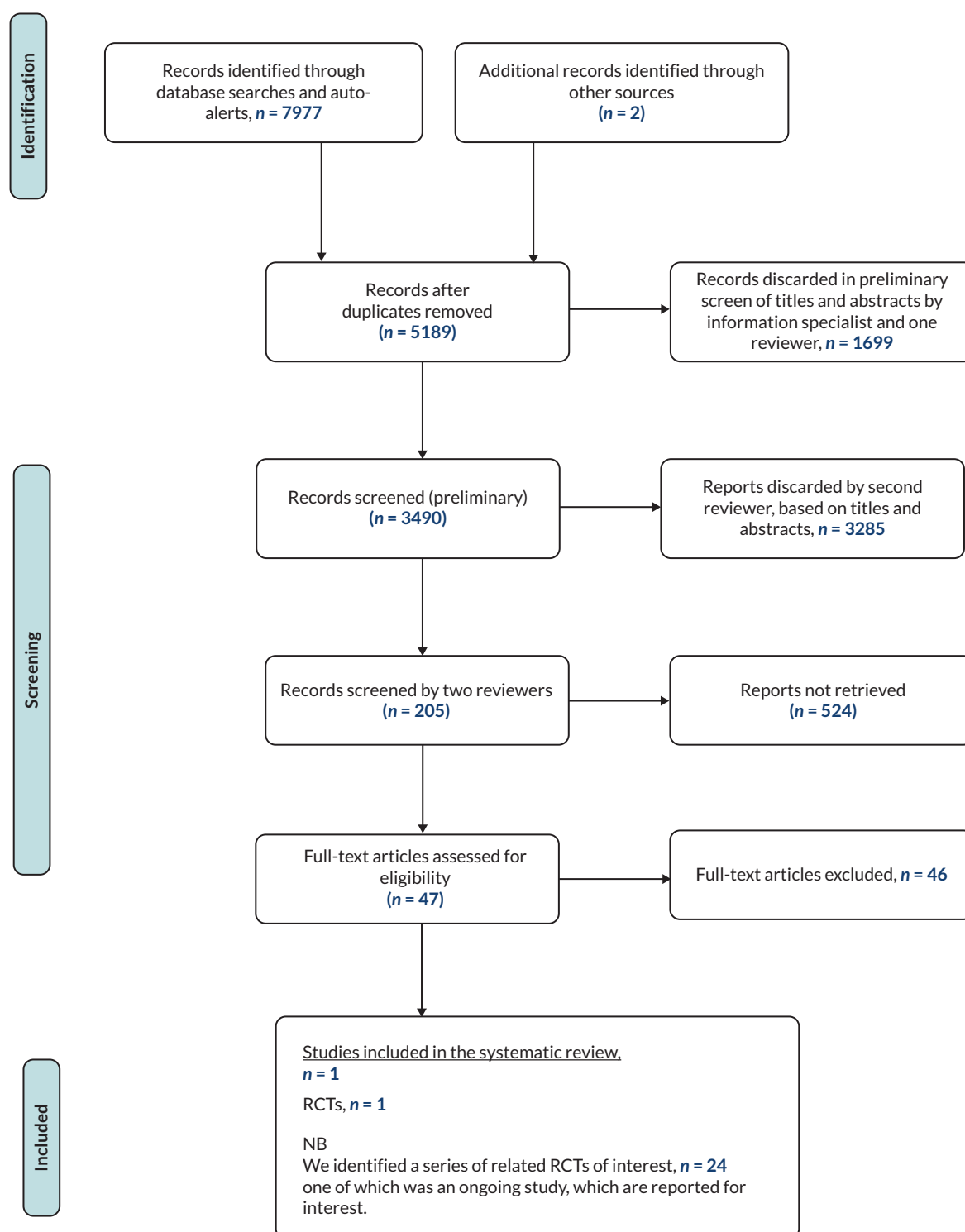


FIGURE 5 PRISMA diagram: biological reconstruction surgery for treatment of early knee OA.

biological surgical reconstruction procedure and type 3: comparisons of load-modifying surgical procedures versus another surgical options.

Type 1: comparisons of biological surgical reconstruction versus non-operative management for treatment of patients with early knee osteoarthritis

There were no studies that compared biological surgical reconstruction to non-operative management in patients with early OA or focal cartilage defects.

TABLE 11 Summary of included biological reconstruction surgery studies characteristics for treatment of early knee OA

Study, design country, quality	Study size and duration	Characteristics: age, % male, BMI, duration of symptoms	Indication and stage OA	Intervention	Comparator	Outcomes
<i>Type 1: comparisons of biological surgical repair surgery in the knee vs. non-operative management to treat early knee OA (RCTs and NRSs)</i>						
No studies						
<i>Type 2: comparisons of biological surgical repair surgery in the knee vs. another biological repair surgery to treat early knee OA (RCTs)</i>						
Kim <i>et al.</i> ¹³² RCT South Korea	89 patients (44 microfracture, 45 AMIC) 2 years' follow-up	Mean age 51.8 (MFx) 48.9 (AMIC) 21% male BMI 25.0	Primary OA 89.9% KL1 18.7% KL2 42.5% KL3 32.5% KL4 6.3% Traumatic OA 10.1%	AMIC 50% with simultane- ous HTO	MF 50% with simultane- ous HTO	VAS KOOS MRI
<i>Type 2: comparisons of biological surgical repair surgery in the knee vs. another type of surgery to treat early knee OA (RCTs)</i>						

Type 2: comparisons of biological surgical reconstruction versus another biological surgical reconstruction procedure for treatment of patients with early knee osteoarthritis

In the Kim study, the authors found that for both treatments there was an improvement in the clinical scores from pre-operation to post operation, seen with KOOS, VAS and IKDC. However, there was no statistically significant difference in scores seen between groups using KOOS (MD 1.9) or IKDC (MD -0.9). There was a MD of -6.0 for the VAS outcome, but again this was not statistically significant.¹³² In summary, the study did not show any significant clinical outcome difference between the groups at 2 years.

Type 3: comparisons of biological surgical reconstruction procedures versus another surgical options for treatment of patients with early knee osteoarthritis

There were no studies that compared biological surgical reconstruction to another surgical strategy in patients with early OA.

Additional analysis

During the search we identified a series of RCTs where biological reconstruction surgery was used to treat cartilage defects (see [Table 12](#)). Nearly all use the presence of OA as a contraindication to their use and none could be used to look at the indirect treatment of early OA. The surgical techniques that are used in these studies can be summarised as follows:

- Mosaicplasty
- MF
- AMIC
- AMIC with bone marrow cells from elsewhere
- ACI
- OCAs

There is significant evidence that relates to the use of ACI in treating isolated cartilage lesions and this is reflected in the NICE technology appraisal guidance for use of this technique.¹⁵⁶ The guidance states that the treatments can only be used if there is no more than minimal OA damage to the knee. This evidence has previously been published as an evidence synthesis and so no further discussion of it is presented in this report.⁴⁴

TABLE 12 Summary of additional comparative studies of biological interventions in the knee

Study, design country, quality	Study size and duration	Characteristics: age, % male, BMI, duration of symptoms	Indication and stage OA	Intervention	Comparator	Outcomes and results
<i>All studies were RCTs comparing on one form of biological surgical repair vs. another form of biological repair in patients with symptomatic cartilage defects. The treatment was not undertaken to prevent development of OA.</i>						
Gudas <i>et al.</i> ¹³⁴ Gudas <i>et al.</i> ¹³³ RCT Lithuania Reports of same study	102 patients 3-year FU Medial condyle	Mean age 32 55% male BMI NR Duration NR but no difference among groups	All having ACL reconstruction. ICRS grades 3–4 KL not reported	Debridement or mosaicplasty	MF	IKDC Tegner
Ulstein <i>et al.</i> ¹³⁵ RCT	25 patients FU 9.8 years (median) Range 4.9–11.4)	Mean age 32.3 56% male BMI not recorded	Cartilage defect (2–6 cm ²) Radiological OA was an exclusion criteria	Mosaicplasty	MF	Lysholm KOOS
Lim <i>et al.</i> ¹³⁶ RCT (three-arm comparison) Korea	30 knees (MF) 22 knees (mosaicplasty) 18 knees (ACI) FU 3 years	Mean age 28.5 years (range 18–42) 58% male	Cartilage defect (mean size 2.8 cm ²)	MF mosaicplasty ACI	Three-arm comparison	Lysholm Tegner HSS MRI
Solheim <i>et al.</i> ¹³⁷ RCT Single-centre Norway	40 patients	Mean age 32 years (range 18–48) 70% male BMI	Cartilage defects (median size 3.5 cm ²) Radiological cartilage narrowing < 4 mm was an exclusion criteria but KL grade not recorded	Mosaicplasty	MF	Lysholm
Kon <i>et al.</i> ¹³⁸ RCT Multicentre Multiple countries	100 patients 2-year FU	Mean age 35 71% and 63% male BMI 25 Previous surgery 50%	Outerbridge 3–4. Lesion size 2–9 cm ² KL NR	MaioRegen scaffold	MF lesions < 4 cm ² Drilling for larger lesions.	IKDC Tegner KOOS VAS
Anders <i>et al.</i> ¹³⁹ Volz <i>et al.</i> ¹⁴⁰ RCT Two centres Germany Two reports of same study	47 patients MF AMIC sutured AMIC glued 5-year FU	Mean age 37 AMIC, 40 MF 77% male BMI 27 and 25	No structural OA at baseline. ICRS 3 or 4	AMIC	MF	Modified Cincinnati ICRS pain score
Shive <i>et al.</i> ¹⁴¹ RCT Canada	62 5-year FU	Mean age 34 AMIC, 40 MF 65% and 54% male BMI 28 and 26	KL grade NR	AMIC (BST CarGel)	MF	WOMAC

Study, design country, quality	Study size and duration	Characteristics: age, % male, BMI, duration of symptoms	Indication and stage OA	Intervention	Comparator	Outcomes and results
Fossum <i>et al.</i> ¹⁴² RCT Norway	41 2-year FU	Mean age 38 Males 40% AMIC, 67% ACI BMI 28 and 26	No structural OA Cartilage defects > 2 cm ²	AMIC	ACI-C	KOOS Lysholm
Knutsen <i>et al.</i> ¹⁴³ Knutsen <i>et al.</i> ¹⁴⁴ Knutsen <i>et al.</i> ¹⁴⁵ RCT Reports of same study	80 2-year FU 5-year FU 15-year FU	Mean age 32 Males 60% BMI not reported	No structural OA Cartilage defects between 2 and 10 cm ²	ACI (periosteum)	MF	Lysholm SF-36
Saris <i>et al.</i> ¹⁴⁶ Van Asche <i>et al.</i> ¹⁴⁷ Vanluawe <i>et al.</i> ¹⁴⁸ RCT Reports of same study	118 1-year FU 2-year FU 5-year FU	Mean age 34 Males 64% BMI not recorded	No structural OA Cartilage defects between 1 and 5 cm ²	ACI (ChondroSelect)	MF	KOOS
Saris <i>et al.</i> ¹⁴⁹ Brittberg <i>et al.</i> ¹⁵⁰ RCT Reports of same study	128 2-year FU 5-year FU	Mean age 34 Males 41% BMI not recorded	Cartilage defects > 3 cm ²	ACI (MACI)	MF	KOOS
Crawford <i>et al.</i> ¹⁵¹ RCT	30 2-year FU	Mean age 40 Male % 83 BMI 28	Cartilage defects ICRS grade 3 Total lesion size < 7 cm ²	ACI (NeoCart)	MF	IKDC KOOS VAS
Basad <i>et al.</i> ¹⁵² RCT	60 2-year FU	Mean age 34 Male 71% BMI 26	Cartilage defects Lesion size 4–10 cm ²	ACI (MACI)	MF	Lysholm
Niemeyer <i>et al.</i> ¹⁵⁴ RCT	102 (97 in final analysis) FU 3 years	Mean age 36, range 18–50. Males 60%. Mean BMI 26	Cartilage defects ICRS grades 3–4 Defects 1–4 cm ²	ACI (Spherrox)	MF	KOOS IKDC
Clave <i>et al.</i> ¹⁵³ RCT France	55 knees 2-year FU	Mean age 29 68% male ACI, 80% mosaicplasty 53% previous surgery	Cartilage defects No structural OA ICRS grades 3–4	ACI (Cartipatch)	Mosaicplasty	IKDC
ACI-C, Autologous chondrocyte implantation-collagen cap; FU, follow-up; HOS-SSS, Hip Outcome Score Sport-Specific Subscale.						

The other studies are recorded here to highlight the surgical procedures that are used to treat patients with cartilage defects. However, a detailed analysis of the studies is outside the scope of this work.

Conclusions

There is no evidence to suggest that the current raft of biological reconstruction techniques available are effective in treating early OA of the knee, when compared to non-operative measures. The only two techniques that have been used in this role, reported in a single study, are MF and AMIC. Both techniques showed an improvement in clinical scores from baseline but there was no comparison between a non-operative treatment programme, and as such this evidence must be considered weak.

As a result of the lack of either randomised controlled studies or high-quality non-randomised controlled studies, comparing biological reconstruction surgery with non-operative treatments there was no evidence that could be taken forward to the health economic analysis.

At present, there is no evidence to support the use of biological reconstruction surgery to treat early knee OA. There is a clear need for more research in this specific area. Due to its effect in treating isolated cartilage defects, a trial of ACI may be the best candidate for treating early OA, but at present NICE does not recommend its use in these patients. Therefore, further research may focus on this potential treatment but only in well-structured studies with feasibility and pilot work leading to a RCT.

Chapter 8 Biological surgical reconstruction for early hip osteoarthritis

Results of the search

The results of the search are described in [Figure 6](#).

The search and successive updates up to January 2021 generated 4293 records and we also identified 32 further studies from other sources. After deduplication and first assessment based on title and abstract screening, we obtained 21 reports in full text. The search strategy is recorded in [Appendix 1](#).

For the intervention review, no RCTs or NRS were found that compared biological reconstruction surgery with non-operative management in patients with early hip OA. In addition, no RCT studies were found that directly compared biological reconstruction surgery with another form of biological reconstruction surgery or another form of surgery in patients with early hip OA. In addition we did not identify any ongoing studies.

Due to the lack of direct comparative studies, we assessed 10 reports of studies that did not meet the criteria of our search (2 reports were of the same cohort).¹⁵⁷⁻¹⁶⁶ However, this material did offer some insight into the use of biological reconstruction surgery in the treatment of patients with early hip arthritis.

Included studies characteristics

The main features of the characteristics of included studies are summarised in [Table 13](#). There were no studies that could be included in the systematic review given the criteria we employed. There were a very small number of NRS that compared one type of biological surgery to another. None of these were randomised controlled studies; therefore, they could not be included in the review and therefore they could not contribute to the health economic analysis.

Additional studies of interest included

Studies included

Due to the lack of studies identified in the review, we have included the most informative of the NRS for addition analysis ([Table 14](#)). These studies did not make the criteria for main review as they are NRS with no comparative data of a surgical intervention against non-operative treatment. Results from these studies cannot be taken forward to the health economic analysis. However, the studies are of interest as they highlight treatments of interest being used in this field.

We identified four NRS that compared the outcome of hip arthroscopy with or without MF in patients with articular cartilage defects and FAI.^{159,161,162,164} A further six NRS are reported that compared the outcome of patients with FAI treated with hip arthroscopy together with two different types of biological reconstruction surgery.^{157,158,160,163,165,166} All patients enrolled had FAI together with some degree of articular cartilage damage. Fontana and Girolamo report on the same series^{158,160} so we will describe nine studies, including the Fontana study. In the 8 cohort studies assessed there were 953 participants. Three studies were prospective cohort studies and five were retrospective studies.¹⁵⁷⁻¹⁶⁶ Study sizes were relatively small, with a range of cohorts from 60 to 140 participants. Five studies were conducted in the USA, two in Italy, one in Canada and one in Spain.¹⁵⁷⁻¹⁶⁶

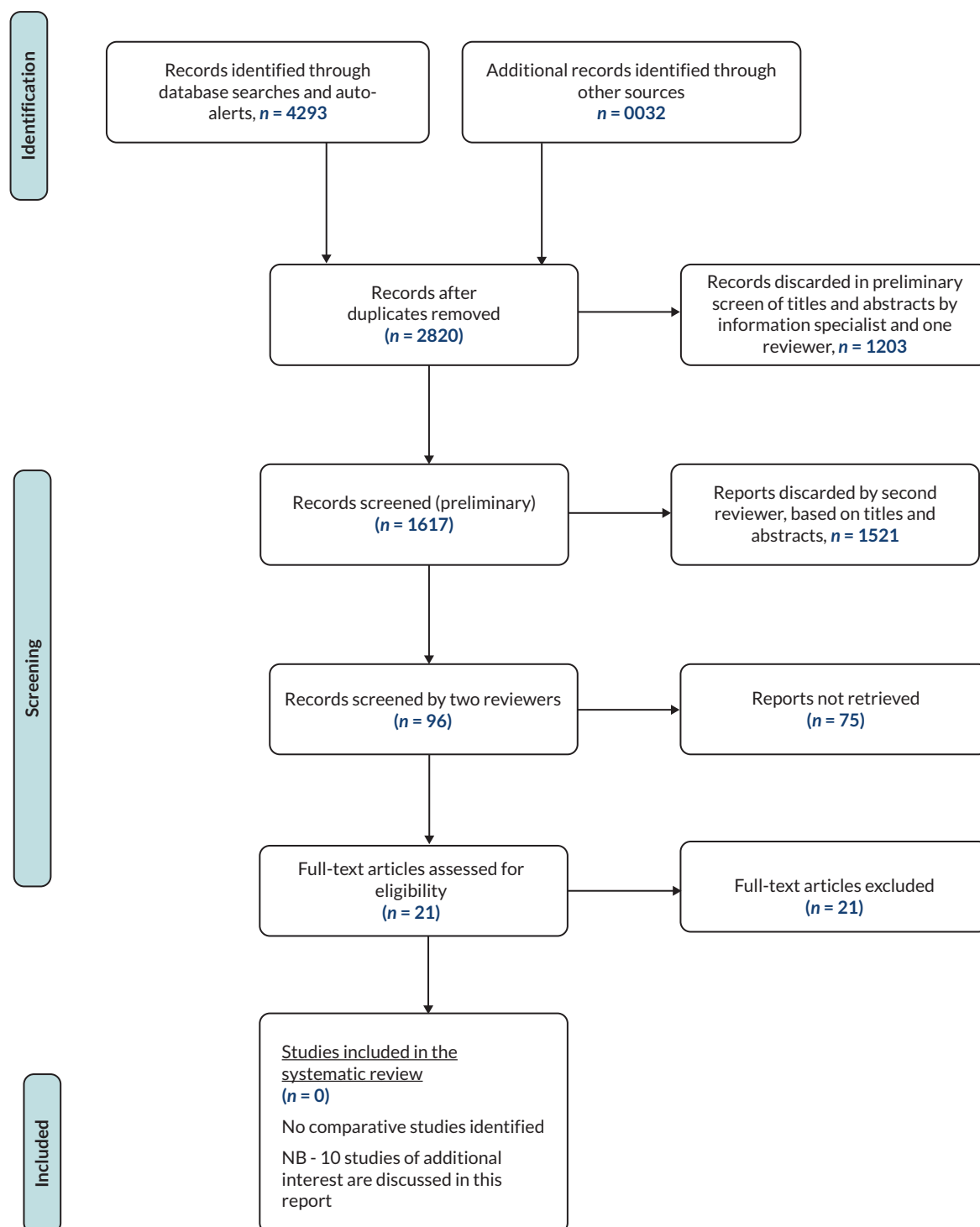


FIGURE 6 PRISMA diagram: biological reconstruction surgery for treatment of early hip OA.

Outcomes

The nine studies used an assessment of pain and function as their primary outcome measure. From the nine studies we assessed, eight studies used the HHS as the primary outcome measure.^{157,159-162,164-166} John *et al.* used the International Hip Outcome Score (iHOT).¹⁶³

Five studies reported outcome at 2 years.^{157,159,161,163,166} Three studies reported at 4–5 years.^{162,164,165} Two studies reported results at medium (3–5 years) and long-term follow-up (8 years) in the same cohort.^{158,160}

TABLE 13 Summary of included biological reconstruction surgery studies characteristics for treatment of early hip OA

Study, design country, quality	Study size and duration	Characteristics: age, % male, BMI, duration of symptoms	Indication and stage OA	Intervention	Comparator	Outcomes and results
<i>Type 1: comparisons of biological surgical repair surgery in the hip vs. non-operative management to treat early hip OA (RCTs and NRSs)</i>						
No studies						
<i>Type 2: comparisons of biological surgical repair surgery in the hip vs. another biological repair surgery to treat early hip OA (RCTs)</i>						
No studies						
<i>Type 2: comparisons of biological surgical repair surgery in the hip vs. another type of surgery to treat early hip OA (RCTs)</i>						
No studies						

Results

Microfracture studies

We reviewed a series of five cohort studies where patients were undergoing arthroscopic non-regenerative surgical treatment for FAI who concurrently underwent MF. Each study compares these patients to a cohort of patients undergoing arthroscopic non-regenerative surgical treatment for FAI who did not undergo MF.

Domb *et al.* reported on patients undergoing surgery for labral tear or FAI, who either underwent MF if they had full-thickness chondral defects or no MF if they did not.¹⁵⁹ The results showed no differences between the MF groups and the controls. Patients in the Hartigan study had femoral MF, mainly when having labral repairs.¹⁶¹ They had grade 4 Outerbridge defects and were compared with a control group with Outerbridge grade 3 or less, and who in the authors' opinion, did not need MF. No significant differences in outcomes were found. It is likely that 9 of the 15 MF patients were also in the Domb study.¹⁵⁹

The Hevesi study compared results in people having labral repairs who had acetabular chondral lesions, treated by MF or debridement and abrasion.¹⁶² Debridement was preferred for defects with substantial surrounding cartilage damage, and for locations where MF was difficult. More of the MF group (36%) had grade 4 defects than the debridement group (9%). No differences in outcomes were found.

The Chaharbakhshi study was again in patients having arthroscopy for labral tears and FAI, and who had acetabular MF.¹⁵⁷ Larger chondral defects (LCDs) were defined as $\geq 200 \text{ mm}^2$ (mean 250, range 200–450) and smaller (SCDs) as $\leq 150 \text{ mm}^2$ (mean 117, range 50–150 mm^2). Both groups showed good improvements and outcomes were similar.

Kester *et al.* report a series of patients who underwent MF in the hip (24 acetabular, 9 femoral) with a control group of 46, matched for age, sex, side, BMI, Tönnis grade and Outerbridge grade 3 or 4, who did not have MF.¹⁶⁴ The authors do not explain why some patients had MF and some did not. Follow-up was at 2 and 4 years. There were no differences in outcomes. mHHSs were improved at 2 years, from baselines 52 to 87 for MF and 84 for controls, and at 4 years, 87 and 77. However, the authors concluded that there were no significant differences between the groups, except for a higher proportion in the MF group requiring hip arthroplasty or resurfacing (18% vs. 2%, $p = 0.038$ but only six patients vs. one).

Comparison of autologous matrix-induced chondrogenesis and microfracture

Fontana and Girolamo provide a non-randomised comparison of AMIC and MF in patients with FAI, with 70 patients having AMIC and 77 having MF.¹⁶⁰ The allocation was determined by whether funding of the collagen matrix was approved by the health authority. Both procedures were carried out in the years 2006–11. The baseline characteristics were fairly similar. Mean defect sizes were similar at 3.7 and 3.5 cm^2 but a slightly higher proportion of the MF group had defect sizes $> 4 \text{ cm}^2$ – 45% versus 39% (not statistically significant). The proportion of men was 71% in the MF group and 51% in the AMIC group. Seventy in each group were followed up at 3 years, and 55% of the MF group and 79% of the AMIC reached 5-year follow-up. Both groups had improved mHHSs to 12 months, but at later time points,

TABLE 14 Summary of additional comparative studies of biological interventions in the hip

Study, design country	Study size and duration	Characteristics; age, % male, BMI	Indication and stage OA	Intervention	Vomparator	Outcomes and results
MF						
Domb 2015 ¹⁵⁹ Non-randomised comparative study USA – American Hip Institute	79 MF, 158 comparison group Minimum 2-year FU	Mean age 48. 59% male. BMI 27	FAI with some patients with full-thickness articular cartilage defects	Hip arthroscopy and MF for full-thickness cartilage defects	Hip arthroscopy and no MF for patients without full-thickness defects	HHS (baseline to 2 years) MF 61 > 78 Comp 60 > 82 HOS-SSS MF 40 > 63 Comp 38 > 69 VAS MF 5.8 > 3.6 Comp 6.0 > 2.8
Hevesi 2019 ¹⁶² Non-randomised comparative USA – American Hip Institute	113 hips Minimum FU 2 years, mean 4.9	60% male, BMI 26 and 28, 33 debridement, 39 MF	FAI Acetabular chondromalacia grade 4 (exposed bone) 8.5% comp group and 35.5% MF	Hip arthroscopy and labral repair and MF in 31. Choice of intervention partly surgeon preference, partly whether location was suitable for MF	Hip arthroscopy and labral repair and Debridement in 82, preferred when there was 'substantial surrounding cartilage damage'	HHS Deb 65 > 86 MF 66 > 81 HOS-SSS Deb 47 > 73 MF 46 > 66 VAS Deb 5.6 > 2.0 MF 5.1 > 2.6 Revision 16% both groups
Hartigan ¹⁶¹ NRS USA – American Hip Institute	15 MF, 45 matched non-MF (it appears 9 were in earlier Domb study)	Mean age 45, 53% male, BMI 26	FAI Outerbridge grade 4 for MF group, 3 or less for controls	Hip arthroscopy and MF femoral head	Hip arthroscopy and MF not considered necessary by surgeon	HHS MF 55 > 72 Comp 63 > 76 HOS-SSS MF 27 > 57 Comp 42 > 68 VAS MF 6.3 > 3.5 Comp 5.2 > 2.6.
Kester 2020 ¹⁶⁴ NRS USA	33 MF, 46 comparison group. 4-year outcomes	Mean age 41, 64% male MF group, 48% comparison group	FAI Outerbridge grade 3 or 4	Hip arthroscopy 33 with acetabular MF and 9 with additional femoral head MF	Hip arthroscopy with no MF 46 hips	HHS MF 52–84 Comp 53 > 78 NAHS MF 49 > 85 Comp 51 > 82
Chaharbakhshi 2019 ¹⁵⁷ NRS USA	107 2-year minimum FU	Mean age 42 (LCD), 43 (SCD) 66% male BMI 27 (LCD) 28 (SCD)	FAI/labral tears LCDs were defined as $\geq 200 \text{ mm}^2$ (mean 250, range 200–450) and smaller (SCDs) as $\leq 150 \text{ mm}^2$ (mean 117, range 50–150 mm^2).	Hip arthroscopy and MF acetabulum	Hip arthroscopy and MF acetabulum	HHS NAHS HOS-SSS
AMIC vs. MF						
Fontana and De Girolamo ¹⁶⁰ De Girolamo ¹⁵⁸ two reports – both NRS Italy	140 patients 3-year minimum FU, most to 5 years Second paper gives results at 8 years in 59 AMIC and 50 MF	Males 71% MF, 51% AMIC ($p = 0.017$)	FAI Outerbridge grade 3 or 4	Hip arthroscopy and AMIC 70	Hip arthroscopy and MF 70	HHS MF 45 > 72 AMIC 45 > 78 Some figures from graphs Early results similar but decline in MF group. By 8 years 11 (22%) of MF group had had THA vs. none in the AMIC group

TABLE 14 Summary of additional comparative studies of biological interventions in the hip (*continued*)

Study, design country	Study size and duration	Characteristics; age, % male, BMI	Indication and stage OA	Intervention	Comparator	Outcomes and results
John 2020 ¹⁶³ NRS Canada	80 patients minimum 2-year FU	70% male. Mean ages 44 MF 36 AMIC BMIs 28 and 27	FAI with labral tear repair or reconstruction	Hip arthroscopy with MF/AMIC (CarGel)	Hip arthroscopy with MF alone	iHOT MF 43–60 CarGel 41 > 58 HOS ADL MF 62–77 CarGel 45–79 THA 35% MF vs. 6% CarGel
AMIC vs. MACI						
Mancini and Fontana 2014 ¹⁶⁵ NRS Italy	57 patients, allocation depending on funding	Males 46% MACI and 42% AMIC, mean ages 36 and 34. Mean defect sizes 2.8 and 2.9 cm ²	for FAI and ace- tabular chondral defects found	Hip arthroscopy with AMIC [Chondro-Gide® (Geistlich Pharma AG, Wolhusen, Switzerland)]	Hip arthroscopy with MACI (BioSed-C)	HHS AMIC 44.9 > 85.5 at 3 years and 80 at 5 years. MACI 46.5 > 85. At 3 years and 83 at 5 years
BMC						
Rivera 2019 ¹⁶⁶ NRS Spain	80 patients, 2-year FU	Mean age 42. 68% male. Outerbridge scores 0–4 but 70% 3 or 4.	FAI	Hip arthroscopy and BMC	Hip arthroscopy with no BMC	HHS Comp 51 > 75 BMC 58 > 87 iHOT Comp 37 > 78 BMC 42 > 86 VAS Comp 6 > 2.7 BMC 5.9 > 1.2
FU, follow-up; LCD, larger chondral defect; SCD, smaller chondral defect.						

the score declined in the MF group but was maintained in the AMIC group, with scores of 72 for MF (from text) and about 78 (from graph). The difference in scores was greater for people with larger defects. The difference in those with small (< 4 cm²) defects is reported as being statistically significant but may not be clinically meaningful, since (from graph, no data in text) it looks close to the MCID of 8 points. Fontana and colleagues conclude that MF alone is the treatment of choice for lesions < 2 cm². During follow-up, six of the MF group required THA, but none of the AMIC group.

Longer-term results from this study are provided by De Girolamo *et al.*¹⁵⁸ The gains in the AMIC group seen at 5 years were mostly maintained to 8 years, with a gain of 37 points in the mHHS. The mHHS scores in the MF group were lower after 2 years and fell further by 5 years, but then fell no further, giving a gain of 28 points. The difference between AMIC and MF at 8 years is only just above the MCID. However, as before, the differences were much less in the group with smaller defects than in those with defects > 4 cm².

In a retrospective study, John *et al.* compared results of hip arthroscopic treatment in 26 patients who had MF alone versus 54 who had MF plus the CarGel scaffold (a gel-forming chitosan-based polymer scaffold – Smith and Nephew, Watford, UK).¹⁶³ Outcome measures used included the international Hip Outcomes Tool-33 (HOT 33) and the HOS – ADL and HOS – Sports Profile. Both groups showed improvement at 2 years, greater in the CarGel group, but not significantly different after adjustment for variables such as age and labral surgery. However, the proportions progressing to THA were 35% (9/26) in the MF alone group and 6% (3/51) in the MF + CarGel group ($p = 0.001$). The mean duration of follow-up was 4 years.

Comparison of autologous matrix-induced chondrogenesis and matrix-associated autologous chondrocyte implantation

Mancini and Fontana provide a non-randomised comparison of AMIC and MACI. AMIC was done using Chondro-Gide, but MACI was done using BioSeed-C (BioTissue, Freiburg, Germany) which has chondrocytes implanted in the Chondral Tissue Scaffold.¹⁶⁵ Allocation to groups was described as 'historical' and made on 'an economic basis' presumably related to availability of funding. Both procedures were carried out by the same person. The baseline characteristics of the 26 MACI and 31 AMIC patients were similar. Mean defect sizes were 2.8 cm² and 2.9 cm². Both groups had good improvements in mHHS scores sustained at 5 years, of 38 and 39 points. Subgroup analysis showed similar results in those with the largest defects (3–4 cm²). So, at 5-year follow-up, AMIC appears as effective as MACI but is much less expensive and more convenient to patients. Longer-term data are needed.

Bone marrow aspirate use in patients undergoing non-regenerative hip surgery for femoro-acetabular impingement

Rivera *et al.* compared results in 40 patients who had surgery for FAI in 2015, and who received BMC, with 40 similar cases from 2104 who did not receive BMC.¹⁶⁶ The cells came from the iliac crest. There were no differences in clinical scores in early follow-up, but at 24 months, the BMC cohort reported less pain and better function. For example, mHHS scores 75 in the control group versus 87 in the BMC group which may not be a clinically meaningful difference.

Conclusions

There are no studies that have directly compared the outcome of biological repair surgery with non-operative measures in patients with early OA. There are no RCT comparisons of different biological repair surgeries or of biological repair compared to other surgery.

We identified some studies that record the use of biological surgery techniques to augment non-regenerative arthroscopic treatment for patients with FAI. These techniques include; MF, AMIC, MACI and bone marrow aspirate (BMA). Within these studies the evidence provided was of very low quality so any interpretation must be made with caution. However, the trend is for no significant clinical advantage for patients with FAI undergoing non-regenerative arthroscopic treatment to undergo an additional biological repair technique.

It is clear that more studies are required to determine if current biological treatments are effective in treating FAI or early hip arthritis. Any trial needs to be a direct comparison of the surgical technique chosen (e.g. MF, AMIC, ACI or stem cell use) with non-operative treatment measures.

Chapter 9 Joint replacement used to treat early knee osteoarthritis

Results of the search

The search and successive updates up to January 2021 generated 12,553 records and we identified 42 other studies through other sources. After deduplication and first assessment based on title and abstract screening, we obtained 211 reports in full text, comprising 32 systematic reviews of interventions and 2 Cochrane review protocols and 128 reports of potential RCTs or comparative NRS ([Figure 7](#)). We also obtained 49 reports of potential prognostic factor studies and 1 systematic review of prognostic factors. We excluded 114 intervention studies and 30 prognosis studies and assessed 35 systematic reviews and protocols, none of which could be used alone to answer the review questions. The search strategy is recorded in [Appendix 1](#).

We included only two studies, both of which were RCTs.^{112,167} The Skou *et al.* study was a trial of TKA versus non-operative care in patients who has early OA of the knee, where a proportion of patients had early OA.¹⁶⁷ The Shaofei *et al.* study has been reported in [Chapter 5](#) and is a RCT of UKA versus HTO in patients with early knee OA.¹¹²

There were no included ongoing studies that would address our systemic reviews question.

In addition, due to the lack of direct comparative studies we included 12 cohort prognostic studies that did not meet the criteria of our search. However, this material did offer some insight into the use of knee arthroplasty in the treatment of patients with early knee arthritis.^{18,168–178}

There were very limited numbers of studies relating to small area resurfacing prostheses in the knee. Their current indication is for focal articular cartilage lesions and we found no studies where the clinical indication is clearly defined early knee OA. Therefore, no studies met the criteria for inclusion in the study.

Included studies characteristics

The main features of the characteristics of included studies are summarised in [Table 15](#).

Study details

We included two RCTs.^{112,167} In addition, due to the lack of primary evidence we included 12 prognostic studies which reported on the outcome of knee arthroplasty performed in patients with early knee arthritis.^{18,168–178} We analysed these separately.

Characteristics of included studies

The included studies are presented in [Tables 15](#) and [16](#). One RCT compared TKA and non-operative management, in 100 patients. It was a two-site study conducted in Denmark and was funded through a foundation.¹⁶⁷ The second RCT compared UKA with HTO in 72 patients. It was a single-site study conducted in China and was funded through public funds.¹¹² The 12 prognostic studies are described in [Table 17](#).

Population, intervention, control/comparison, outcome characteristics

Summary details of the PICO characteristics are given in [Tables 15](#) and [16](#).

Surgical interventions assessed

We report two types of surgical intervention: joint replacement surgical intervention for people with isolated lateral or medial compartment early OA disease and joint replacement surgical intervention for people with early OA other than isolated lateral or medial compartment disease. We note that some studies may have defined their population

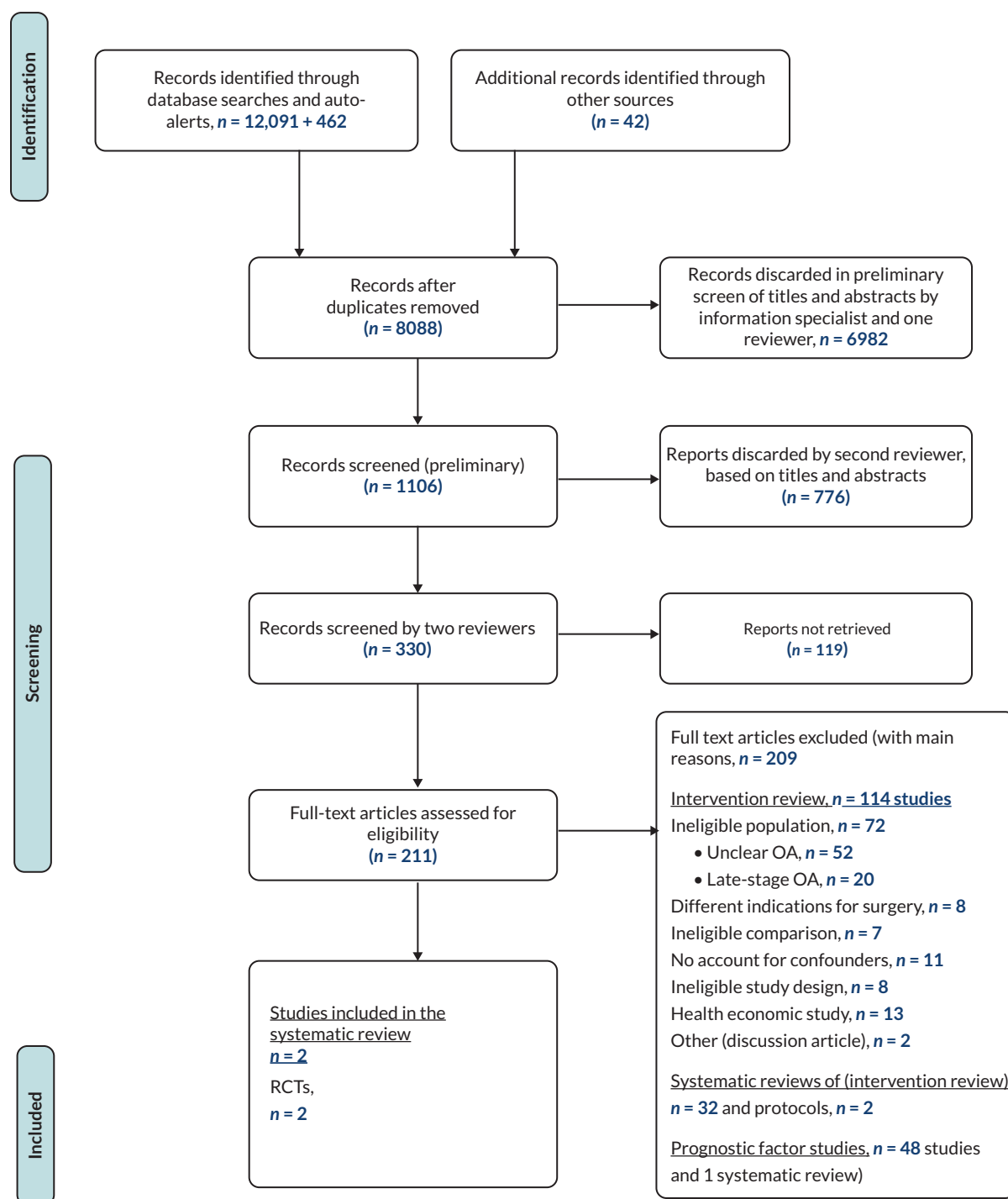


FIGURE 7 PRISMA diagram: knee arthroplasty surgery for treatment of early knee.

as 'suitable for TKA', but this does not necessarily mean bi- or tri-compartmental disease – some surgeons may not perform UKAs, so would classify people with unicompartmental disease as suitable for TKA.

The following surgical interventions, stratified as above, were reported in the included studies. Some studies allocated participants to a combination of surgery and non-operative comparator (e.g. TKA plus non-operative management).

- TKA – one RCT.¹⁶⁷
- UKA (medial or lateral) – one RCT.¹¹²

TABLE 15 Summary of included joint arthroplasty surgery studies characteristics for direct treatment of early knee OA

Study, design and country	Study size and duration	Population characteristics: mean age (SD), sex; male (%) BMI (kg/m ²)	Indication and stage of OA	Intervention	Comparator	Adjustment for confounding factors (method/factor adjusted for)	Outcomes
<i>Type 1: joint replacement surgical procedures vs. non-operative management in early knee OA (RCTs or NRS)</i>							
No studies identified							
<i>Type 2: joint replacement vs. another joint replacement procedure in patients with early knee arthritis (RCT)</i>							
No studies identified							
<i>Type 3: joint replacement surgical procedures vs. another type of surgery in patients with early knee arthritis (RCT)</i>							
Shaofei et al. ¹¹² RCT (single-centre) China	72 (36, 36) 2 years	<ul style="list-style-type: none"> Age: 53.9 (7.4) and 52.5 (8.4) Sex: 3/36 (8%) and 5/36 (14%) BMI: NR 	<ul style="list-style-type: none"> HTO group: KL2: 10/36 (28%) KL3: 26/36 (72%) UKA group: KL2: 11/36 (31%) KL3: 25/36 (69%) 	HTO	UKA	RCT	Combined pain and function, pain
Medial compartment							
NR, not reported.							

TABLE 16 Summary of included joint arthroplasty surgery studies characteristics for indirect treatment of early knee OA

Study, design and country	Study size and duration	Population characteristics: mean age (SD), sex; male (%), BMI (kg/m ²)	Indication and stage of OA	Intervention	Comparator	Adjustment for confounding factors (method/factor adjusted for)	Outcomes
Type 1: joint replacement surgical procedures vs. non-operative management to indirectly treat patients with early knee OA (RCTs or NRS)							
<i>RCTs</i>							
Skou et al. ¹⁶⁷ (2 centres) Denmark	100 (50, 50) 2 years	<ul style="list-style-type: none"> Age: 65.8 (8.7) and 67.0 (8.7) Sex: 18 (36%) and 20 (40%) BMI: 32.3 (6.2) and 32.0 (5.8) Unclear compartment 	<ul style="list-style-type: none"> TKA group: KL1: 0%, TKA KL2: 7/50 (14%), KL3: 21 (42%) KL4: 22/50 (44%) Non-operative management group: KL1: 0%, KL2: 5/50 (10%), KL3: 21 (42%) KL4: 24/40 (48%) 		Non-operative management (education, exercise/physical therapy, analgesia and weight loss plus insoles)	RCT	Combined pain and function, pain, function, HRQoL, AEs/SAEs, reoperation/revision
Type 2: joint replacement vs. another joint replacement procedure to indirectly treat patients with early knee arthritis (RCT)							
No studies identified							
Type 3: joint replacement surgical procedures vs. another type of surgery to indirectly treat patients with early knee arthritis (RCT)							
No studies identified							

Comparator interventions assessed

The following non-operative comparators were reported in the included studies:

- Non-operative management: combination of exercise, education, dietary advice, use of insoles and pain medication (12 weeks) – one RCT.¹⁶⁷
- HTO – one RCT.¹¹²

Outcomes

All studies reported at least one major outcome described in our protocol. Most outcomes related to pain or physical function, combined pain and function, and HRQoL used at least one validated instrument. Both studies only reported results at short-term follow-up (2 years). The following outcomes were reported:

- Pain: two RCTs.^{112,167}
- Physical function: one RCT.¹⁶⁷
- Pain and function: two RCTs.^{112,167}
- HRQoL: one RCT.¹⁶⁷
- Surgical complications and AEs: two RCTs.^{112,167}

Results

We present the results in three main sections, type 1: comparisons of joint replacement procedures versus non-operative management; type 2: comparisons of one joint replacement versus a different type of joint replacement and type 3: comparisons of one joint replacement versus another surgical procedure.

Direct treatment of early knee osteoarthritis

Type 1: comparison of joint replacement surgical procedures versus non-operative management used to directly treat people with early OA.

There are no studies to report.

Type 2: comparisons of joint replacement versus another type of joint replacement used to directly treat people with early knee OA.

There are no studies to report.

Type 3: comparisons of joint replacement versus another surgery procedure used to directly treat people with early knee OA.

One RCT of UKA versus HTO in 72 participants reported PROMs outcomes at 2 years' follow-up. All studies were in knees with medial compartment OA.¹¹² This has been reported in [Chapter 5](#). The evidence is of very low certainty. RCT evidence from 1 study in 72 participants suggests there may be a moderate benefit for UKA compared to HTO at 2 years' follow-up for pain, and for combined pain and function. For example, for pain and function (HSS knee score, 0–100), the MD was 10.3 (6.4–14.2), favouring UKA, and the change from baseline for HTO was 28 HSS points. At long-term follow-up, the NRS findings were inconclusive for PROMs outcomes and both NRS had features that made the findings unreliable. Very low certainty evidence for a meta-analysis of the results from these two studies in 350 participants for the reoperation outcome suggests the hazard for HTO may be about twice that for UKA for conversion to TKA at long-term follow-up (up to 30 years) – HR = 2.13 (95% CI 0.79 to 5.56), with some heterogeneity in the results. This corresponds to 115 more people per 1000 who had received HTO (rather than UKA) having revisions to TKA, from 53 fewer to 185 more.

Indirect treatment of early knee osteoarthritis

Type 1: comparison of joint replacement surgical procedures versus non-operative management used to indirectly treat people with early OA.

There are no studies that directly compare joint replacement surgical procedures versus non-operative management in people with early OA. One RCT in 100 participants compared TKA with non-operative management, but not exclusively in patients with early OA.¹⁶⁷ Almost half the participants in the study had late-stage OA (KL grade 4 (KL4): 44% TKA and 48% non-operative management). We have not been able to obtain subgroup data for early OA from the authors, and therefore report these results as indirect evidence. Very low certainty RCT evidence (downgraded for high risk of bias, indirectness and imprecision) suggests that there may be a large benefit for TKA compared to non-operative management at 12 months' follow-up for each of pain, physical function and QoL. For example, the MD for pain (KOOS pain score, 0–100) was –16.7 (95% CI –24.1 to –9.35), favouring TKA; the change from baseline for the non-operative management group was 17 KOOS points. Very low certainty evidence from the RCT suggests considerably more subsequent operations for the non-operative management group, such that 32% of patients randomised to the non-operative group switched to the TKA group during the follow-up period, while only one in the TKA group had revision surgery.

Type 2: comparisons of joint replacement versus another type of joint replacement used to indirectly treat people with early knee OA.

There are no studies in this group.

Type 3: comparisons of joint replacement versus another surgery procedure used to indirectly treat people with early knee OA.

There are no studies in this group.

Additional analysis

We assessed 12 prognostic studies that all examined outcomes following knee replacement surgery in people who had OA, comparing patients with early OA disease to those with late OA pre-operatively (see [Table 17](#)). All studies were aimed at determining the prognostic value of early OA, independent of the effect of age, sex, BMI and preoperative scores. We note that the mean ages of the participants in the 12 prognosis studies were between 62 years and 70.8 years; thus, the adjustment for age does not extend to younger people (under 55 years) receiving joint replacements.

We included studies in which early OA was defined as in our protocol KL1–3, but we also examined studies that defined early OA as KL1–2 or partial-thickness OA to probe this boundary for joint replacement. As expected, these studies report large changes from preoperative values for PROMs outcomes for people with late OA following joint replacement (e.g. an improvement of 30 WOMAC pain points (0–100 scale) for people receiving TKA, and 17 OKS pain and function points (0–48 scale) for UKA). However, any differences between early and late OA were small: for example, 2.9 (0.13–5.72) WOMAC pain points and 2.6 (95% CI 0.19 to 5.0) OKS points for the same analyses. Thus, there are some benefits of joint replacement in people with early OA.

Three cohort studies assessed the outcome of UKA in 429 participants (433 knees).^{169,172,174} They all defined early OA in terms of cartilage loss, reporting the association of partial-thickness cartilage loss (PTCL) versus FTCL with PROMs and reoperation/revision outcomes. All studies preoperatively matched patient characteristics between groups of PTCL and FTCL, one of which used propensity score matching. The evidence is mainly of very low certainty and suggests that associations with PROMs outcomes may be either small – with PTCL (early OA) giving poorer outcomes – or that there is little difference between early and late OA. For example, for physical function, it is possible that there is no important difference between early and late OA at short-term and medium-term follow-up [e.g. a mean difference of 1.80 AKSS function points on a scale of 0–100 (95% CI –2.48 to 6.08)] at 12 months, but there may be a small association of PTCL with worse physical function at long-term follow-up (MD 6.10, 0.97–11.2) on the same AKSS scale. However, this is very low certainty evidence. In low certainty evidence for the combined pain and function outcome at short-term follow-up, meta-analysis of the results from two studies in 320 participants suggests a small to moderate association of PTCL with worse combined pain and function (MD 4.24, 0.46–8.01 for the OKS, 0–48), with unimportant heterogeneity ($I^2 = 68\%$) – in the same direction of association. However, at longer term, very low certainty evidence from a larger study at 5 years' follow-up suggests early OA may be associated with slightly more pain and function than late OA – MD 2.60 (0.19–5.01) on an OKS scale of 0–48. Across the three studies (one at short-term follow-up and two at long-term follow-up), few participants had reoperations: 30 in 433 participants, of which only 8 had conversions to TKAs (from UKAs); consequently, the evidence is very uncertain. Very low certainty evidence from a meta-analysis of three studies for the outcome, any reoperation, including conversion (of UKA) to TKA, suggests that reoperation at up to 5-years' follow-up may be more likely following UKA in people with PTCL compared with FTCL (RR 3.28, 1.59–64). This corresponds to 98 more people per 1000 having reoperations in the early OA group, from 25 to 247 more, for a FTCL group risk of 43 people per 1000. For the outcome, conversion to TKA, 8 in 346 people in 2 studies had TKA conversion in 2 studies. Very low certainty evidence suggests that, at long-term follow-up, 58 more people per 1000 with early OA may have a TKA conversion (from 3 fewer to 357 more) for a late OA group risk of 18 people per 1000 (RR = 4.22, 0.85–20.9).

Nine studies assessed the outcome of TKR in 2953 participants reported the association of early OA (vs. late OA) with PROMs outcomes.^{18,168,170,171,173,175–178} The majority of studies took into account key confounding factors using multivariable regression analyses. There is methodological and clinical heterogeneity among the studies, for example, different outcome scales and cut-off values for dichotomous outcomes, different durations of follow-up, different cut-off points for the prognostic factors and the reporting of change scores/final values and adjustment for confounders. Consequently, we did not conduct many meta-analyses and report a narrative summary here.

TABLE 17 Summary of additional prognostic studies of interest where joint arthroplasty surgery has been used in the treatment of early knee OA

Study, design and country	Study size and duration	Population characteristics: mean age (SD), sex; male (%) BMI (kg/m ²)	Indication and stage of OA	Intervention	Comparator	Adjustment for confounding factors (method/factor adjusted for)	Outcomes
Cushnaghan <i>et al.</i> ¹⁶⁸ Prospective follow-up of cases in a case-control study (3 centres) UK	657 (KL4: 188, KL0–2: 137, KL3: 286) 6 years (mean, range 2–8 years) from TKA	<ul style="list-style-type: none"> KL4: 96 (32%) KL0–2: 62 (20%) KL3: 144 (48%) NB. Numbers analysed; 23 participants did not have KL values	<ul style="list-style-type: none"> TKA wait-listed patients Age: < 70: 123/259 (47%); 70–75: 67 (26%); > 75: 69 (27%) Sex: 93/259 (36%) BMI: < 25: 37/259 (14%); 25 to < 30: 134 (52%); ≥ 30: 88 (34%) (overall population) 	Primary OA. Number with multicompartamental disease NR.		Multivariable regression (age, sex, BMI, smoking, hypertension, diabetes, thyroid disease, previous injury to knee, Heberden's nodes, number of painful joint sites in hands, shoulders, and knees)	Function
Dowsey <i>et al.</i> ¹⁸ Prospective cohort study (single-centre) Australia	501 (KL < 3 : 15 modified KL3a: 57, KL3b: 200, KL4a: 87, KL4b: 119) 12 months NB. 23 not included due to missing radiographs	<ul style="list-style-type: none"> KL < 3: 15 (3%) KL3a: 57 (12%) KL3b: 200 (42%) KL4a: 87 (18%) KL4b: 119 (25%) (modified KL grades) 	<ul style="list-style-type: none"> Primary TKA patients Age: 70.8 (8.3) Sex: 147 (30.8%) BMI: 32.2 (6.0) (overall population values) 	Uni- and multicompart-ment OA	<ul style="list-style-type: none"> Unicompartment 32% (medial 22%, lateral 8%, PF 2%, unclear 1%) Multicompartment 67% (medial and PF 51%, lateral and PF 8%, medial and lateral 1%, tricompartment 7%) 	Multivariable regression [age, sex, BMI, age adjusted Charlson Comorbidity Index (CCI), baseline pain and function scores, SF-12 mental and physical function scores]. Surgical variables: patella resurfacing and type of prosthesis	Pain, function
Hamilton <i>et al.</i> ¹⁷⁴ Retrospective cohort study (single-centre) UK	278 (282 knees): PTCL: 90 (94 knees) FTCL: 188 (188 knees) 5 years	<ul style="list-style-type: none"> PTCL: 94 (33%) FTCL: 188 (67%) (KL grade NR) 	<ul style="list-style-type: none"> Patients with UKA (requested following failure of other treatments) Age: 64.2 (11) and 63.6 (10) Sex: 40 (43%) and 84 (45%) BMI: 29.1 (5) and 29.7 (5) 	Medial compartment OA		Propensity score matching 1 : 2 (age, sex, pre-operative OKS)	Pain and function, revision, reoperation, survival, progression
Study, design and country	Study size and duration	Prognostic factor categories/cut points	Population characteristics mean age (SD), sex; male (%) BMI (kg/m ²)	Indication for surgery		Adjustment for confounding factors (method/factor adjusted for)	Outcomes
Hoorntje <i>et al.</i> ¹⁷⁵ Retrospective cohort study (single-centre) The Netherlands	327 (KL3–4: 144, KL 1–2: 47) (numbers analysed) 2.3 years (median) range: 1.9–2.6 years	<ul style="list-style-type: none"> KL3–4: 144 [KL3: 117 (61%); KL4: 27 (14%)] KL1–2: 47 [KL1: 3 (2%); KL2: 44 (23%)] 	<ul style="list-style-type: none"> Primary TKA patients Age: 68.8 (8.4) and 70.2 (9.3) Sex: 59 (41%) and 16 (34%) BMI: 29.9 (5.0) and 28.4 (3.8) 	Primary OA. Number with multicompartamental disease NR.		Multivariable regression (age, sex, BMI)	Pain, function

continued

TABLE 17 Summary of additional prognostic studies of interest where joint arthroplasty surgery has been used in the treatment of early knee OA (*continued*)

Study, design and country	Study size and duration	Population characteristics: mean age (SD), sex; male (%) BMI (kg/m ²)	Indication and stage of OA	Intervention	Comparator	Adjustment for confounding factors (method/factor adjusted for)	Outcomes
Keurentjes <i>et al.</i> ¹⁷⁰ cohort study (4 centres) The Netherlands	277 (KL0–2: 22, KL3–4: 191 (90%) (numbers analysed) 2.82 years (mean)	<ul style="list-style-type: none"> KL0–2: 22 (10%) KL3–4: 191 (90%) (numbers for each KL grade NR) 	<ul style="list-style-type: none"> Primary and revision TKA patients Age: 65.1 (10.3) and 69.5 (8.6) Sex: 31.8% and 30.9% BMI: < 25: 33.3%, 25–30: 27.8%, 30–35: 33.3%, > 35: 5.6% and < 25: 14.7%, 25–30: 46.7%, 30–35: 21.2%, > 35: 17.4% 	OA. Number with multicompartmental disease NR.		Multivariable regression (age, sex, BMI, and Charnley classification)	Pain, function, PREMs
Lange <i>et al.</i> ¹⁷⁷ RCT analysed as a prospective cohort study (single-centre) USA	304 [OARSI < 10: 102, OARSI grade ≥ 10: 138 (numbers analysed) 2 years	<ul style="list-style-type: none"> OARSI grade < 10: 102 (42%) OARSI grade ≥ 10: 138 (58%) 	<ul style="list-style-type: none"> Primary TKA patients Age: 66.8 (8.5) and 66.4 (7.7) Sex: 36 (35.3%) and 57 (41.3%) BMI: 30.2 (6.1) and 30.72 (5.78) 	Primary OA. Number with multicompartmental disease NR		Multivariable regression (age, sex, BMI, comorbidity score) NB. From RCT data	Pain, function
Maier <i>et al.</i> ¹⁷² Retrospective cohort study (single-centre) Germany	64 (PTCL: 32, FTCL: 32) PTCL group: 3.8 years (mean) FTCL group: 4.0 years (mean)	<ul style="list-style-type: none"> PTCL: 32 (50%) FTCL: 32 (50%) 	<ul style="list-style-type: none"> UKA patients Age: 62 (8.0) and 63 (7.8) Sex: 15 (47%) and 15 (47%) BMI: NR 	Anteromedial OA		Simple matching (age, sex, pre-operative OKS, time to follow-up)	Combined pain and function, pain, function, revision
Pandit <i>et al.</i> ¹⁶⁹ Retrospective cohort study (single-centre) UK	87 [PTCL: 29, FTCL (no bone loss): 29], FTCL (bone loss): 29) 2 years (mean)	<ul style="list-style-type: none"> PTCL; normal to PTCL: 29 (33%) FTCL (no bone loss): 29 (33%) FTCL (with bone loss): 29 (33%) 	<ul style="list-style-type: none"> UKA patients Age: 62 (10.4), 60 (8.5) and 63 (11) Sex: NR BMI: NR 	Medial compartment OA		Simple matching (age, sex, pre-operative OKS, and follow-up): matching of two different groups (FTCL (bone-on-bone), with or without bone loss) to PTCL group	Combined pain and function, re-operations
Rehman <i>et al.</i> ¹⁷⁸ Prospective cohort study (single-centre) Norway	156 (KL4: 38, KL < 4: 118) 12 months	<ul style="list-style-type: none"> KL4: 38 (24%) KL < 4: KL2, 3 (2%), KL3, 115 (74%) 	<ul style="list-style-type: none"> TKA patients Age: 69 (9) and 69 (9) Sex: 50 (32%) and 50 (32%) BMI: 29.2 (4.6) and 29.2 (4.6) 	Primary multicompartment OA. Unicompartment OA excluded		Multivariable regression (age, sex, BMI, preoperative pain level, comorbidities prior ipsilateral knee surgery, prior contralateral TKA). NB. Results for covariates only for QoL	Pain, function

Study, design and country	Study size and duration	Population characteristics: mean age (SD), sex; male (%) BMI (kg/m ²)	Indication and stage of OA	Intervention	Comparator	Adjustment for confounding factors (method/factor adjusted for)	Outcomes
Riis <i>et al.</i> ¹⁷¹ Prospective cohort study (single-centre) Denmark	176 (KL4: 131, KL0–3: 45) 12 months	<ul style="list-style-type: none"> KL4: 131 (74%) KL0–3: KL0, 1 (1%), KL1, 10 (6%), KL2: 11 (6%), KL3, 23 (13%) 	<ul style="list-style-type: none"> TKA patients Age: 68 (range 39–91) and 68 (range 39–91) Sex: 63 (36%) and 63 (36%) BMI: ≥ 30 kg/m²: 73 (41%) and ≥ 30 kg/m²: 73 (41%) 	Primary OA. Number with multicompartamental disease NR.		Multivariable regression (age, sex, BMI, pre-operative pain, drainage, haemoglobin decrease)	Pain, function
van der Water <i>et al.</i> ¹⁷⁶ Prospective cohort study (single-centre) The Netherlands	559 (KL0–3: 486, KL4: 73) 2 years	<ul style="list-style-type: none"> KL0–1: 66 (12%) KL2: 85 (15%) KL3: 335 (60%) KL4: 73 (13%) 	<ul style="list-style-type: none"> TKA patients Age: KL0–1: 64 (10), KL2: 67 (9), KL3: 67 (10), KL4: 67 (9) Sex: KL 0–1: 14 (21%), KL2: 29 (34%), KL3: 119 (36%), KL4: 19 (26%) BMI: KL0–1: 29 (4), KL2: 29 (4), KL 3: 29 (5), KL4: 29 (5) 	Primary OA. Number with multicompartamental disease NR.		Baseline similarities for categorical prognostic factor across KL grades for BMI and preoperative pain, but borderline differences for age and sex. [Multivariable analysis only for continuous prognostic factors (adjusted for BMI, age, sex, SF-12 MCS, pre-operative pain)]	Pain, function
Tilbury <i>et al.</i> ¹⁷³ Prospective cohort study (single-centre) The Netherlands	417 (KL 0–2: 74, KL3–4: 197) 1 year	<ul style="list-style-type: none"> KL 0–2: 74 (27%) KL 3–4: 197 (73%) 	<ul style="list-style-type: none"> TKA patients Age: KL1–2: 67 (10) years; 0–55 years: 6 (8%), 56–65 year: 23 (31%), 66–75 years: 31 (42%), 76–85+: 14 (19%) Age: KL3–4: 67 (10) years; 0–55 years: 22 (11%), 56–65 years: 61 (31%), 66–75 years: 74 (38%), 76–85+: 40 (20%) Sex: KL1–2: 18/74 (24%) and KL3–4 62/197 (31 %) BMI: 30 (4.5) and 30 (4.7) 	Primary OA. Number with multicompartamental OA NR.		Multivariable regression (age, sex, BMI and pre-operative scores of PROMs assumed)	Combined pain and function pain, function, HR-QoL
PF, patellofemoral							

We assessed studies where we could compare early OA defined KL < 4 compared to late OA defined as KL4. There is low certainty evidence from four studies, at short-term follow-up only, reporting the association of KL < 4 versus KL4 with pain or with physical function suggests that associations with PROMs outcomes may be either small, with KL < 4 giving poorer outcomes, or that there is little difference between early and late OA. All the studies had limitations, such as (1) an indirect outcome (pain and discomfort); (2) an indirect prognostic factor (OARSI score, which had a cut-off point that meant 16% participants with 'early OA' had KL4 knees); (3) an unadjusted analysis; or (4) a dichotomised outcome (with unclear dichotomising thresholds). There was very low certainty evidence for pain as a dichotomous outcome at short-term follow-up. For function, low certainty evidence suggests no important difference in EQ-5D mobility subscale: SMD = 0.16 (-0.31 to 0.52), and meta-analysis of unadjusted KOOS ADL results from 2 studies in 600 people suggests a possible small association of KL < 4 with worse function at short-term follow-up: MD 3.74 (0.71–6.76) (scale 0–100). One study in 154 people, reporting function as a dichotomous outcome – low (0–80) versus high (81–100) on a KSS scale – suggests there may be poorer function outcomes for people with KL < 4: OR = 1.48 (0.67–3.27), but the wide CI demonstrates the lack of certainty in the findings. Two studies (one with unadjusted analyses) analysed the different KL grades as categorical variables, allowing us to investigate any 'dose effects' according to KL grade. Both sets of results were low certainty evidence: the unadjusted analysis suggests KL0–1 and KL2 (in comparison with KL4) may each be associated with more pain and with poorer function, whereas KL3 had little association (similar results to those for KL4). For example, KL2 versus KL4: MD 12.5 (2.6–22.4) KOOS pain points (scale 0–100), but KL3 versus KL4: MD 1.0 (-5.3 to 7.3). The second study investigated the prognostic value of modified KL grades in comparison with KL4b (which was defined as KL4 plus any subchondral bone attrition, making this indirect evidence). The evidence suggests a 'dose effect' for pain: compared with KL4b, KL3a was associated with more pain, but KL4a showed no difference in pain. The corresponding results for function were very low certainty evidence.

We assessed studies where we could compare early OA defined as KL0–2 with late OA defined as KL3–4. At medium-term follow-up, for pain and for function, low certainty evidence from two studies in 191 and 213 participants, suggests KL0–2 may be associated with more pain and poorer function than KL3–4. Methodological differences precluded meta-analysis, but there were similar findings for the two studies: for example, *for pain*, mean differences of 5.7 (-2.9 to 14.3) and 11.0 (4.4–17.6) KOOS pain points; and, *for function*, mean differences of 6.4 (-2.4 to 15.2) KOOS ADL points and 19.1 (8.5–29.7) SF-36 physical functioning points. At long-term follow-up (6–7 years), there may be slightly poorer function for KL0–2 compared with each of KL3 and KL4: KL0–2 versus KL3 – MD = 7.6 (-0.2 to 15.3) SF-36 physical function points and KL0–2 versus KL4 had an MD of 5.6 (-2.8 to 14.0) on the SF-36 physical function scale (0–100). However, we note that this long-term evidence had many missing data (ca. 60%) and is of very low certainty. One study in 448 participants reported the independent association of multicompartment OA (vs. unicompartment OA) with each of pain and function at 2 years post TKA: moderate certainty evidence shows a small association with reduced pain for multicompartment OA, reporting a MD of -3.1 (-6.3 to -0.3) IKSS pain points (scale 0–50) and no difference for function: MD = -1.34 (-6.00 to 3.32).

Conclusions

No studies directly compared joint replacement surgical procedures with non-operative management solely in people with early OA and therefore no data to take forward to the HE analysis.

There was indirect evidence from one RCT studying TKA versus knee OA, that included a population with about 50% early OA. Very low certainty evidence suggests there may be a large benefit in terms of pain, function and QoL for joint replacement surgical procedures in comparison with non-operative management, and considerably more reoperations under RCT conditions for the non-operative management group compared with TKA.

There is also very limited addition indirect evidence from the comparison of arthroplasty versus other surgical procedures.

Our assessment of prognostic studies suggests patients with early knee OA can see clinical benefit after joint replacement, but there may be more pain after surgery in for patients with early OA (KL < 4) compared with late OA (KL4), suggesting a worse outcome for these patients. However, this is all based on weak evidence.

To address the question of whether knee arthroplasty (partial or total) is an effective treatment for early OA, there would need to be an appropriately sized RCT in patients with only early OA comparing joint replacement surgery versus non-operative management with short-, medium- and long-term follow-up. Care must be taken in recruiting patients to any such study given the risk of poor outcome, for some patients noted in the prognostic review.

Chapter 10 Joint replacement performed for early hip osteoarthritis

Results of the search

The search and successive updates up to January 2021 generated 12,806 records and we also identified 32 further studies from other sources ([Figure 8](#)). After deduplication and first assessment based on title and abstract screening, we obtained 133 reports in full text, comprising 20 systematic reviews of interventions and 2 Cochrane review protocols and 76 reports of potential RCTs or comparative NRS. We also obtained 31 reports of potential prognostic factor studies and 4 systematic reviews of prognostic factors. The search strategy is recorded in [Appendix 1](#).

For the intervention review, no RCTs or NRS was found that compared joint replacement surgery with non-operative management in patients with early hip OA. In addition, no RCT studies were found that directly compared hip joint replacement surgery with another form of joint replacement surgery or another form of surgery in patients with early hip OA.

We also identified 1 systematic review of 5 NRS in 431 hips comparing THA to PAO in people with hip dysplasia.¹⁷⁹ Two of the studies included in that review were conducted in populations with at least some of the participants having early OA. However, the Tönnis grade differed at baseline between intervention groups in one study, and in the other study, the OA severity was not reported for the THA group at baseline, which may indicate that underlying pathology had an influence on the intervention provided. The remaining studies identified by the review assigned those with more severe OA to THA and the less severe participants to PAO and/or did not report the OA severity. Only one of the studies carried out propensity score matching, and the others did not take account of confounding. Consequently, we have not included this review nor any of the studies in it.

Therefore, we could not include any intervention studies. In addition we did not identify any ongoing studies.

Due to the lack of direct comparative studies, we included six cohort prognostic studies that did not meet the criteria of our search. However, this material did offer some insight into the use of knee arthroplasty in the treatment of patients with early hip arthritis. They were all cohort studies that compared the outcome of hip replacement in early versus late hip OA.^{18,168,170,173,180,181}

Similar to the findings in the knee 9 (see [Chapter 9](#)), there were very limited numbers of studies relating to small area resurfacing prostheses used in the hip. Their current indication is for focal articular cartilage lesions and we found no studies where the clinical indication is clearly defined early hip OA. Therefore, no studies met the criteria for inclusion in the study.

Included studies characteristics

The features of the characteristics of included studies are summarised in [Table 18](#). There were no studies that met our search criteria.

Additional analysis of studies outside the search criteria

Studies included

We identified six studies into our additional analysis that compared the outcome of THR in patients with early or late hip OA ([Table 19](#)). Only one with direct information for the prognostic value of early OA (KL1–3 or equivalent) compared to late OA, as defined in our review protocol; one that used the Croft classification, comparing grade 0–3 versus 4 versus 5; and we also considered four further studies for this review: three studies compared KL1–2 versus KL 3–4;

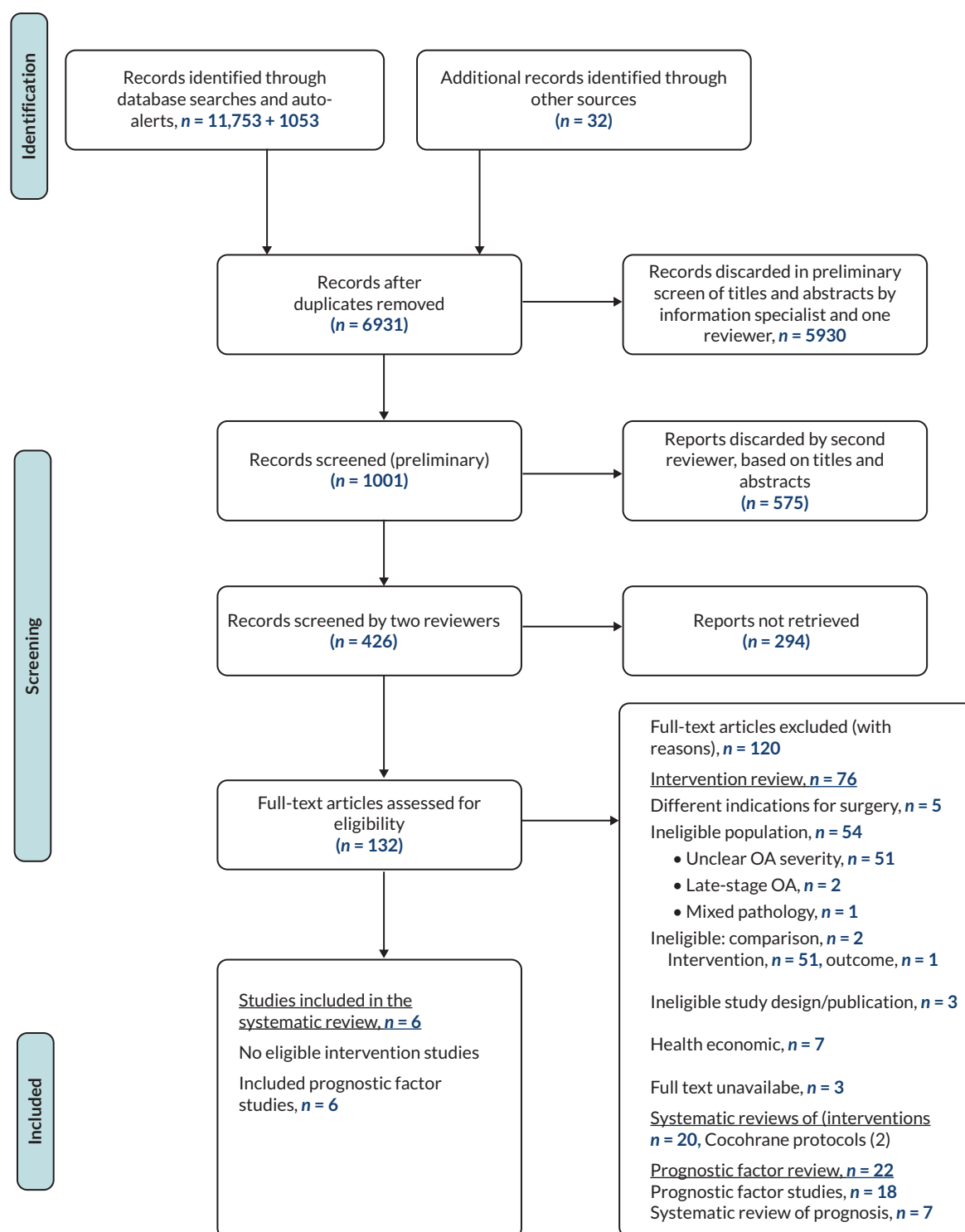


FIGURE 8 PRISMA diagram: knee arthroplasty surgery for treatment of early hip OA.

and one study split severity into six modified KL grades (KL1, KL2, KL3a, KL3b, KL4a, KL4b), combining KL1–3a and comparing this, KL3b and KL4a with KL4b (complete loss of joint space with subchondral bone attrition).^{18,168,170,173,180,182}

The 6 cohort studies involved a total of 3475 enrolled participants. Three were prospective cohort studies; one was a cohort study within a case-control study; one study had two retrospective analyses of a prospectively collected data and one was a RCT which was analysed as a cohort study.^{18,168,170,173,180,182}

TABLE 18 Summary of included joint arthroplasty surgery studies characteristics for treatment of early hip OA

Study ID, design and country	Study size and duration	Population: mean age (SD), sex; male (%), BMI (kg/m ²)	Pathology and indication for surgery (KL/Tönnis grade)	Intervention	Comparator	Adjustment for confounding factors (method/factors adjusted for)	Outcomes
<i>Type 1: comparison of hip arthroplasty vs. non-operative management in patients with early hip OA (RCTs and NRS)</i>							
No studies identified							
<i>Type 2: comparison of hip arthroplasty vs. another type of hip arthroplasty procedure in patients with early hip OA. (RCTs)</i>							
No studies identified							
<i>Type 3: comparison of hip arthroplasty vs. another type of surgery in patients with early hip OA. (RCTs)</i>							
No studies identified							

The Judge *et al.* study reports on a group of patients that were previously reported in 2010, but we only include the 2011 study.¹⁸² Study sizes ranged from 250 to 1327 with a median of 437. Two studies were conducted in the UK.^{168,182} A further two were based in the Netherlands.^{170,173} One study was conducted in Australia and one in Sweden.^{18,180} Three studies had non-commercial support.^{168,173,181} One study had both commercial and non-commercial support for the trial on which the study was based.¹⁷⁰ One study stated no funding and one did not disclose details of financial support.^{18,180}

The six studies reported mean ages between 66 years and 69 years. BMI was reported in all studies, with mean BMI corresponding to overweight (25–29.9 kg/m²). All studies were predominantly in women. All studies included hips with primary OA considered suitable for THA and accounted for key confounding factors (age, sex, BMI) with multivariable regression analyses.

Prognostic factors assessed

The prognostic factor for this review is early OA (vs. late OA), which was variously defined:

- KL < 4 versus KL4: 1 study.¹⁸¹
- Croft 0–3 versus 4 versus 5 (this may correspond to KL0–2 vs. KL3 vs. KL4, but the literature is unclear): 1 study.¹⁶⁸ The report analysed Croft grade as a categorical value; for this study the proportion of Croft grades were: 0–3 = 38 (15%); Croft 4 = 174 (70%), Croft 5 = 37 (15%).
- KL0–2 versus KL3–4: 3 studies.^{170,173,180}
- Modified KL grades (KL2 + 3a, KL3b, KL4a vs. KL4b): 1 study.¹⁸

Outcomes assessed

We prioritised the PROMs outcomes of pain, function, and combined pain and function for the prognosis review. No studies reported revision/reoperation rates. Four studies reported results at short-term follow-up (≤ 2 years).^{18,173,180,181} One study at medium-term follow-up (> 2 years to < 5 years).¹⁷⁰ Only one study reported long-term follow-up.¹⁶⁸

The following outcomes were reported:

- Pain: two studies reported pain as a continuous outcome.^{170,173} One study reported a dichotomous outcome as the number of patients achieving the minimally important difference over baseline.¹⁸ Three studies did not report on pain.^{168,180,181}
- Physical function: three studies reported a continuous outcome.^{168,170,173} A single study reported a dichotomous outcome.¹⁸ Two studies reported the SF-36 physical function subscale.^{168,170} One study reported the HOOS ADL and one reported HHS function.^{18,173} Two studies did not report this outcome.^{180,181}
- Pain and function: two studies reported this outcome as a continuous variable.^{173,180} One study reported pain and function as a dichotomous outcome.¹⁸¹ Two studies used the WOMAC total scale.^{180,181} One study used the OHS.¹⁷³ Three studies did not report this outcome.^{18,168,170}

TABLE 19 Summary of additional prognostic studies of interest where joint arthroplasty surgery has been used in the treatment of early hip OA

Study, design and country	Study size and duration	Prognostic factor categories/cut points	Population characteristics: mean age (SD), sex; male (%) BMI (kg/m ²)	Adjustment for confounding factors (method/factor adjusted for)	Outcomes
Prognostic studies					
Al-Amiry <i>et al.</i> ¹⁸⁰ Prospective cohort study (single-centre) Sweden	250 (73, 149) 12–15 months	<ul style="list-style-type: none"> KL1–2 ('mild'): 71 (32%) KL3–4 ('severe') = 149 (68%) 	<ul style="list-style-type: none"> Primary OA treated with THA Age: KL1–2 = 68.2 (9.8); KL3–4 = 67.8 (9.9) Sex: 30 (41%) and 105 (70%) BMI: 27.6 (4.6) and 27.6 (4.4) 	Multivariable regression (age, sex, BMI, baseline WOMAC)	Combined pain and function, QoL, PREMs
Cushnaghan <i>et al.</i> ¹⁸³	249 (38, 174, 37) mean 8.5 years	<ul style="list-style-type: none"> Croft 0–3 (moderate) = 38 (15%) 	<ul style="list-style-type: none"> Participants 45 years or older with primary hip OA 	Multivariable regression (age, sex, BMI, baseline score, smoking, hypertension, diabetes, previous injury to hip, Heberden's nodes, no. of painful joint sites in hands, knees, shoulders)	QoL, function
Retrospective cohort study Two health districts in England (combined population of ≈1 million residents) UK (first analysis)		<ul style="list-style-type: none"> Croft 4 (severe) = 174 (70%) Croft 5 (very severe) = 37 (15%) <p>NB. original data source (643 patients) had 96% KL3 or 4</p>	<ul style="list-style-type: none"> Age: 67.2 (8.2) Sex: 89 (36%) BMI: 27.3 (4.2) <p>(overall population values)</p>		
Dowsey <i>et al.</i> ¹⁸ Prospective cohort study (single-centre) Australia	382 2 years	<ul style="list-style-type: none"> KL2 and modified KL3a: 30 (8%) 3b: 125 (33%) 4a: 95 (25%) 4b: 132 (34%)—comparison KL2 and modified 3a, 3b and 4a vs. 4b 	<ul style="list-style-type: none"> Age: 68.9 (9.3) Sex: 150 (39%) BMI: 29.9 (5.5) <p>(overall population values)</p>	Multivariable regression (age, sex, BMI, baseline pain and function and SF-12 mental and physical function scores and ASA score, surgical approach, femoral head size, cemented/cementless)	Pain, function
Judge <i>et al.</i> ¹⁸¹ Prospective cohort study (20 centres) UK	1327 12 months	<ul style="list-style-type: none"> KL0–3 (554) vs. KL4 (497) (2011) Age as a continuous variable 	<ul style="list-style-type: none"> Age: 65.7 (10.9) Sex: 559 (42.1%) BMI: 27.5 (4.4) 	Multivariable regression for OA severity as a prognostic factor (age, sex, BMI, preoperative WOMAC, preoperative EQ-5D), number of expectations, education, ASA grade, no. of medications)	Combined pain and function
Keurentjes <i>et al.</i> ¹⁷⁰ Retrospective cohort study on RCT data (4 centres) The Netherlands	445 mean 2.8 years	<ul style="list-style-type: none"> KL0–2: 92 (33%), KL3–4: 185 (67%) (numbers analysed, unclear number at baseline) 	<ul style="list-style-type: none"> Primary orthopaedic patients 18 years and older considered for primary or revision THA for OA Age: 65.1 (7.8) and 67.4 (8.7) Sex: 23.9% and 44.3% BMI: < 25: 29.2%, 25–30: 41.6%, 30–35: 23.6%, > 35: 5.6% and < 25: 34.1%, 25–30: 46.4%, 30–35: 16.2%, > 35: 3.4% 	Multivariable regression (age, sex, BMI, and Charnley classification) NB. from RCT data	Pain, function, PREMs
Tilbury <i>et al.</i> ¹⁷³ Prospective cohort study (single centre) The Netherlands	428 (77, 225) 1 year	<ul style="list-style-type: none"> KL0–2: 77 (25%) KL3–4: 225 (75%) (indirect evidence) 	<ul style="list-style-type: none"> Age: 66 (11) years; 0–55 years: 12 (16%), 56–65 years: 23 (30%), 66–75 years: 29 (38%), 76–85 +: 13 (17%) and 68 (9) years; 0–55 years: 17 (8%), 56–65 years: 71 (32%), 66–75 years: 93 (41%), 76–85 +: 44 (20%) Sex: 32 (42%) and 93/225 (41%) BMI: 28 (4.5) and 27 (4.4) 	Multivariable regression (age, sex, BMI, preoperative scores of PROMs, Charnley classification)	Combined pain and function, QoL, function, QoL

Results

The included evidence reported in this chapter is solely represented by a prognosis review, which examines the independent prognostic value of early OA (vs. late OA) for outcomes following THA. We included studies in which early OA was defined as in our protocol (KL grades 1–3), but also examined other studies that compared KL0–2 with KL3–4 in order to probe the effectiveness boundary for THA, and this comparison formed most of the evidence. The larger volume of evidence for KL0–2 versus KL3–4 may reflect how hip specialists define early OA of the hip. Most of the evidence was for outcomes at short-term follow-up, but one study reported findings at medium-term and another at long-term follow-up.

From the outset, we note that there were large changes from preoperative values for PROMs following THA in people with more severe OA (e.g. an improvement of 46 HOOS pain points (0–100 scale) and 26 SF-36 physical functioning points at 12 months' follow-up for people receiving THA with KL grades 3–4). In comparison, differences between early and late OA were relatively small at short-term and medium-term follow-up, so there may still be some benefits of joint replacement in people with early OA. At long-term follow-up differences in physical function may be large between early and late OA.

In participants receiving THA, 6 studies in 3475 eligible participants (1851 analysed participants) reported the association of early OA (vs. late OA) with pain, function, or combined pain and function. No studies investigated associations with reoperation outcomes. All studies adjusted for key confounding factors using multivariable regression analyses. There is methodological and clinical heterogeneity among the studies: for example, different outcome scales and cut-off values for dichotomous outcomes, different durations of follow-up, different cut-off points for the prognostic factors and the reporting of change scores/final values (which represent different research questions). Consequently, we did not conduct any meta-analyses and report a narrative summary here.

Most of the evidence is of very low certainty, mainly because of risk of bias, some imprecision and indirectness in defining both prognostic factors and outcomes. Generally, there were small to moderate associations of early OA with worse PROMs outcomes or there was no association.

Only one study reported the association of early OA as defined by our protocol (KL grade < 4) versus late OA (KL4), for the pain and function outcome at short-term follow-up.¹⁸¹ Very low certainty evidence suggests that KL < 4 (vs. KL4) may be associated with worse pain and function, but there is uncertainty: OR = 1.30 (95% CI 0.84 to 2.00) for non-responders defined using OMERACT-OARSI criteria based on the total WOMAC score [we note that for prognostic factor studies, a MD greater than zero, or an OR > 1 indicate that early OA is associated with worse outcomes]. A different analysis of the same data (adjusted for different confounders) compared KL2 versus KL4 (OR = 0.83, 0.14 to 5.00) and KL3 versus KL4 (OR 1.67, 0.83–3.33), but there is too much uncertainty to determine if there is a dose response.

Three studies in 888 participants reported PROMs outcomes in patients with KL0–2 versus KL3–4: at short-term follow-up.^{170,173,180} Low/very low certainty evidence from one study in 302 participants suggests a small or moderate association with worsening pain [MD = 6.10 (2.20–10.00) HOOS pain points (scale 0–100)] and poorer physical function [MD = 7.50 (2.80–12.20) HOOS ADL points (scale 0–100)]. Low/very low certainty evidence from two studies in a total of 524 participants suggests there is small or no association with the combined pain and function outcome, for example, OHS (assumed scale of 0–48), MD= 2.0 (–0.2 to 4.2) and 1.00 (–3.29 to 5.29) WOMAC points (0–100).

At medium-term follow-up, however, very low certainty evidence from another study in 277 participants suggests there may be no difference in pain between the two OA severity groups: MD 0.88 (–6.08 to 7.84) SF-36 bodily pain points (scale 0–100). In contrast, for function (SF-36 physical function), there may be poorer outcomes for KL0–2 by 8.93 points (2.14–15.72), scale 0–100.

One study in 249 participants assessed OA severity using the Croft scheme (which may correspond to the KL grades as Croft 0–3 being similar to KL0–2) and investigated independent associations with function.¹⁶⁸ The evidence suggests

there may be a fairly large association: Croft 0–3 was associated with poorer function in relation to each of Croft 4 and Croft 5, which, respectively, had MDs of 8.3 (–0.9 to 17.5) and 19.4 (7.7–31.1).

One study analysed modified KL grades as categorical variables, allowing us to investigate any ‘dose effects’ according to modified KL grade at short-term follow-up: very low certainty evidence suggests large associations with pain for KL3b and KL4a versus KL4b (e.g. OR 7.69, 0.87–68.0), although CIs are wide.¹⁸ There may be a ‘dose effect’ for function: compared with KL4b (which the authors defined as KL4 plus any subchondral bone attrition), KL3b and KL4a were each associated with poorer function (OR 3.85, 1.77–8.33 and OR = 2.70, 1.24–5.88); this is low/moderate certainty evidence.

Conclusions

There are no studies that have directly compared the outcome of hip arthroplasty with non-operative measures in patients with early OA.

The only evidence identified is from prognosis studies comparing independent associations of early OA (vs. late OA) with PROMs outcomes.

From our review hip joint arthroplasty is performed in patients who have early hip OA. The only evidence identified is from prognosis studies comparing independent associations of early OA (vs. late OA) with PROMs outcomes. However, there is weak evidence that the outcome of the intervention in this group is worse than when compared to late-stage hip OA (KL4). This evidence is mainly of very low certainty, but suggests that early OA may be associated with poorer hip function, especially at medium-term and long-term follow-up. Additionally, there appears to be worse pain at short-term follow-up in people with early OA.

Our main finding is that there is no evidence to measure the efficacy of the hip replacement in early hip OA hip arthroplasty when compared to non-operative measures. This lack of evidence is important as patients with early hip OA will face a shared decision-making process around choosing hip replacement or continued non-operative care. It appears that the results of hip replacement in patients with late hip OA cannot be extrapolated to patients with early hip OA.

There is a clear research question to answer. Is hip replacement more effective than non-operative measures in patients with early hip OA?

Chapter 11 Cost-effectiveness of surgical interventions for early knee and hip osteoarthritis

Introduction

The aim of this analysis was to examine the cost-effectiveness of alternative treatment strategies for early-stage structural OA of the knee and the hip. Decision-analytic (cost-utility) models were developed to assess the cost-effectiveness of surgical versus non-operative management procedures. The alternative strategies of the cost-effectiveness analyses were established based on the findings from the systematic review of the available clinical evidence. Of the four groups of surgical interventions, only non-regenerative procedures were found to show potentially sufficient clinical evidence to allow for a modelling of their cost-effectiveness compared to non-surgical treatment.

Evidence synthesis of the clinical effectiveness identified the following surgical procedures to show potentially sufficient evidence for their cost-effectiveness to be assessed:

- For early knee OA: *arthroscopic debridement* (including some meniscal excision) compared to *non-operative management* such as physiotherapy and mixed conservative treatments.
- For early hip OA: *arthroscopic osteochondroplasty* versus *non-operative management* for FAI (morphological problems), which are among the principal causes of early hip OA, compared to *non-operative management* such as physiotherapy and mixed conservative treatments.

This chapter describes the models' structure, the parameters used within the models (transition probabilities, resource use, costs, and utilities), the assumptions made, the different scenarios evaluated results, and the SAs undertaken.

Model structure

Two cohort state-transition 'Markov' models were developed to reflect the clinical pathways of patients with early OA of the knee and the hip. The structure of the models was built over an iterative process of consultation with expert clinicians, surgeons and physiotherapists involved in the study. Although other modelling frameworks could have also been employed (decision tree, discrete event simulation), a Markov model was considered appropriate given that the patient pathway of the population of interest can be well represented in the health states, and the transitions between each health state, which would not be dependent on what patients experienced previously, beyond their initial intervention (surgery or non-operative management). Markov models have also been used in previous economic studies given the chronic nature of the disease and the ability of such models to handle recurrent events.¹⁸⁴

The models were developed over a time horizon of 5 years given the lack of evidence on the long-term costs and outcomes of patients receiving procedures for early OA. Simulated patient cohorts in the Markov model moved through clinically and economically distinct health states. Transitions between states were modelled on a yearly basis meaning that each time the model 'runs', a year of 'model time' has passed. Therefore, terms 'year' or 'cycle' in this chapter are used interchangeably. Each health state, excluding death, had associated costs and quality-adjusted life-years (QALYs) values. Both costs and QALYs were 'triggered' in every cycle the patient was on that health state. The different health states for the model are shown by the rectangles in [Figure 1](#). The arrows indicate the transitions patients can make in the model. The loop arrows in the right-hand corner of the rectangles (recurring arrow) mean that a patient can stay in that health state for more than one cycle, and potentially indefinitely (until they die).

Separate models for each joint (hip and knee) were developed to reflect the fact that different data inputs were found in the systematic review of the literature. The structure of the Markov models of the hip and the knee was assumed identical in terms of time horizon, number and sequence of health states. The analysis was conducted from the perspective of the NHS.

Patients with early OA of the knee or hip have marked symptoms and disability that may not settle with first-line treatments and can be as severe as those with end-stage disease who are considering joint replacement. Following the NICE clinical guideline on OA, patients who present with clinical OA should undergo first-line non-operative treatments, such as good information, physical therapy and weight loss, as well as second-line measures, such as knee injections, if appropriate.¹⁸⁵

The structure of the Markov model was designed to reflect the condition of patients who, after trying first- and second-line treatments, are said to have entered the 'treatment gap', where symptoms and disability can still be significant, and it is not clear which subsequent intervention is most value for money. The condition of patients with early OA who try first-line treatments and still have persistent symptoms is represented by the dotted health states in [Figure 9](#).

At this stage, patients enter the Markov model by receiving either non-operative management or arthroscopic debridement which are the alternative strategies assessed in the cost-effectiveness models. The Markovian health states in which patients receive one of these two interventions are called the 'intervention states'. Patients 'transit' in these health states for just one cycle, after which they move to the 'postintervention' states (post management or post arthroscopy). Patients in the post-intervention states can stay for longer than just a cycle, until their early OA condition worsens suggesting the need of a joint replacement. It is assumed that within the post-management state, there would be patients who would respond favourably to the intervention and those who would not, and, as a result, they would be expected to have different levels of QoL and associated healthcare costs. Although the model makes no distinction between them because there are no data reporting on outcome subgroups following interventions for patients with early OA, the QoL and healthcare costs associated to the post-intervention states reflect what is known about all patients who receive the respective treatments regardless of outcome. However, if the outcome is such that patients merit a joint replacement, then the model does address this by allowing for patients to transit to the primary replacement state. Subsequently, patients move to the post-primary state, where they can stay indefinitely, or they can move out and receive a revision operation. After a revision, patients remain in the post-revision state for the remaining time. We assumed there was no mortality risk from non-operative management. Arthroscopic debridement is reported to be a safe procedure with small risk of complications and negligible risk of surgical mortality; therefore we also assumed there was no risk of surgical death.¹⁸⁶ As a result, patients can move to death from any of the health states due to all-cause mortality. The clinical pathway of patients with early OA is illustrated in the Markov model shown in [Figure 9](#).

Model inputs

Inputs to the Markov models were selected based on the results of the reviews of the existing clinical studies which were assessed for non-regenerative procedures and their comparators for early OA of the knee and the hip.

The model for early knee OA was informed by one RCT study by Kirkley *et al.* and one comparative NRS by Su *et al.* reporting the comparison between arthroscopic debridement (including some meniscal excision) and physiotherapy.^{88,94}

The synthesis of the clinical studies on early OA of the hip found three RCTs and one NRS by Scott *et al.* reporting the comparison of interest (arthroscopic osteochondroplasty vs. non-operative management).^{103-105,107}

In the next section, we describe how the evidence reported by these studies was used to inform the transition probabilities determining the number of patients progressing throughout each Markovian state (see [Transition probabilities](#)), the QALYs (see [Health-related quality of life](#)) and corresponding healthcare costs (see [Resource use and costs](#)) associated with each health state for each comparator.

Transition probabilities

In both knee and hip models, we assumed the probability of patients moving out of the intervention, primary replacement and revision health states to the corresponding post-intervention, post-primary and post-revision states equal to one risk of death. The probability values specific to each model are illustrated in the next sections.

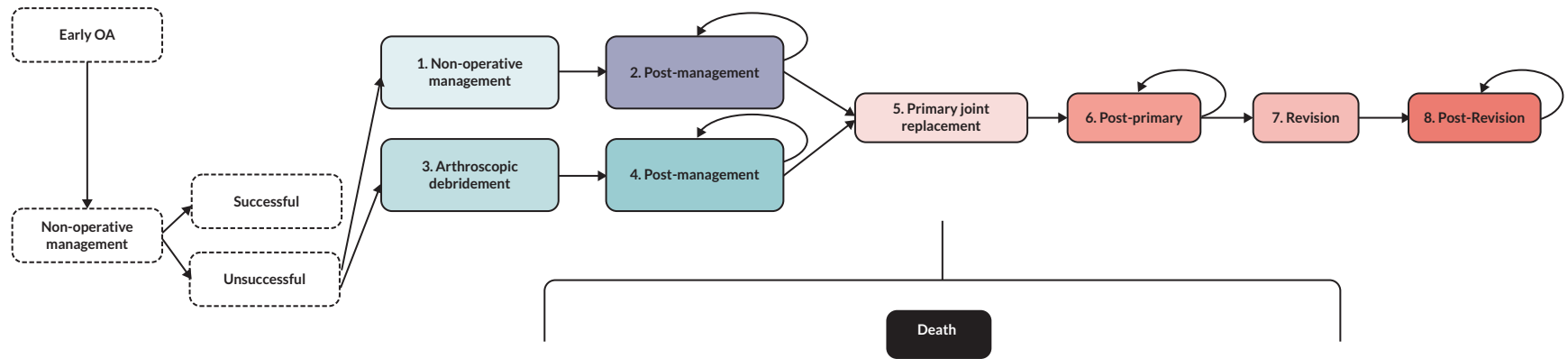


FIGURE 9 Markov model for knee and hip analyses.

Knee: transition to joint replacement

The evidence synthesis suggested there may be fewer TKRs in people who had surgery at 5 years' follow-up for the non-randomised comparison of arthroscopic debridement (with some MF) versus non-operative management.⁹⁴ To create annual transition probabilities for the primary joint replacement health state, time-to-event data from Su *et al.* were digitised and pseudo individual-level data created using the algorithm from Guyot *et al.*^{94,187} Log-cumulative hazards plots were inspected for proportional hazards. Parametric models (exponential, Weibull, Gompertz, lognormal, log-logistic, gamma and generalised gamma) were fitted and assessed for goodness-of-fit via the Akaike information criterion (AIC) and Bayesian information criterion (BIC) and visual inspection over the period of observed data. Extrapolation was not necessary for the cost-effectiveness model and was therefore not assessed. Inspection of log-cumulative hazards plots showed evidence of proportionality; thus, it was deemed appropriate to fit survival models with treatment group as a covariable. The Gompertz distribution was chosen as the most appropriate model as informed by the AIC and BIC values as well as visual inspection, and therefore used in the base-case analysis. The resulting annual probabilities are shown in [Figure 10](#).

Knee: transition to revision

Studies in the evidence synthesis did not report reoperation or revision rate outcomes for the comparison of arthroscopic debridement versus physiotherapy. One study by Gu *et al.* investigated the relationship between prior knee arthroscopy within 2 years of TKR on revision rates after TKR.¹⁸⁸ This study was originally excluded by the evidence team as it was unclear whether the population presented with early OA diagnosis. Given the lack of evidence assessing risk of revision on early OA populations, we decided to employ this study in our model to capture the potential effect of knee arthroscopy on the probability of revision by assuming that when patients undergo TKR they present with severe symptoms due to end-stage structural OA. Annual probabilities of revision were derived from the reported all-cause incidence of revision at 2 years. We converted probabilities covering multiple years into instantaneous event rates and subsequently into annual revision probabilities following established methods.¹⁸⁹ As a result, the model was informed with distinct but constant probabilities of revision for patients for the non-operative versus arthroscopy arms ([Table 20](#)).

Hip: transition to joint replacement ([Table 21](#))

Evidence from one NRS in 846 participants suggested a small increase (not statistically significant) in the number of people having a THR at 2 years' follow-up in the surgical group when compared with conservative care.¹⁰⁷ We recreated Kaplan–Meier data from the study and extrapolated values up to 5 years; however, proportionality did not hold with any of the tested distributions with substantial difference at 5 years, which conflicted with the conclusion of the NRS study. Therefore, probability values for the two treatment arms were obtained by converting the reported rate at 2 years into annual probabilities and by informing the model with constant rates for the two arms.¹⁸⁹

Hip: transition to revision (see [Table 21](#))

Additional evidence comparing the impact of prior arthroscopic surgery (vs. no surgery) on outcomes post THR at short-term and medium-term follow-up showed an increased revision rate for patients who had prior arthroscopy.¹⁹¹ Annual

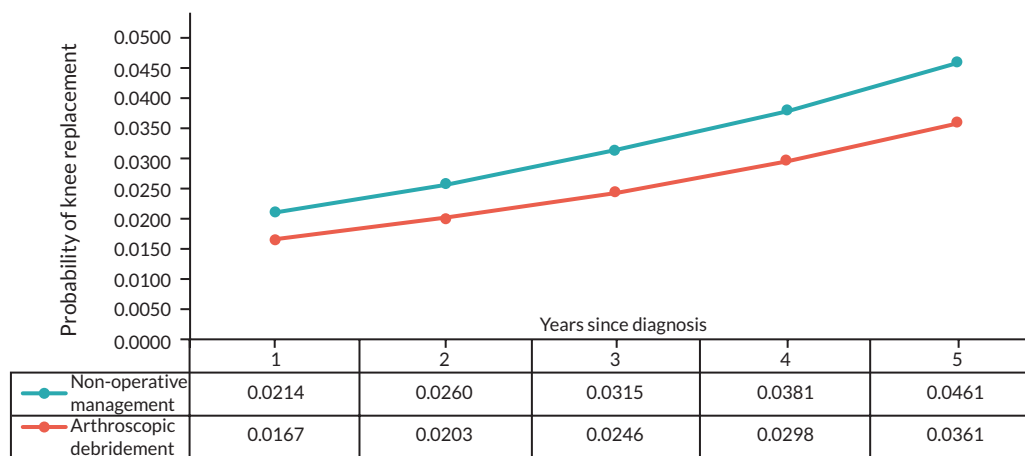


FIGURE 10 Annual transitions to knee joint replacement.

TABLE 20 Transitions probabilities – knee model

Years since diagnosis	Transition to death (59–63 years old)	Transition to primary joint replacement – non-operative management arm	Transition to primary joint replacement – arthroscopic debridement arm	Transition to revision – non-operative management arm	Transition to revision – arthroscopic debridement arm
1	0.0058	0.0214	0.0167	0.0090	0.0139
2	0.0064	0.0260	0.0203	0.0090	0.0139
3	0.0069	0.0315	0.0246	0.0090	0.0139
4	0.0078	0.0381	0.0298	0.0090	0.0139
5	0.0084	0.0461	0.0361	0.0090	0.0139
Source	¹⁹⁰	⁹⁴	⁹⁴	¹⁸⁸	¹⁸⁸

TABLE 21 Transitions probabilities – hip model

Years since diagnosis	Transition to death (35–40 years old)	Transition to primary joint replacement – non-operative management arm	Transition to primary joint replacement – arthroscopic debridement arm	Transition to revision – non-operative management arm	Transition to revision – arthroscopic debridement arm
1	0.0008	0.0197	0.0223	0.0104	0.0171
2	0.0009	0.0197	0.0223	0.0104	0.0171
3	0.0010	0.0197	0.0223	0.0104	0.0171
4	0.0010	0.0197	0.0223	0.0104	0.0171
5	0.0011	0.0197	0.0223	0.0104	0.0171
Source	¹⁹⁰	¹⁰⁷	¹⁰⁷	¹⁹²	¹⁹²

rates of revision were not reported by the available studies. Therefore, we informed our model with the revision rates at 2 years reported by Mahalias *et al.* who had shorter time frame compared to the other available evidence which we deemed more relevant for our model given the fact that revision is the last health state in the model.¹⁹² Revision rates at 2 years were converted into annual constant probabilities for the two groups.¹⁸⁹

Health-related quality of life

According to NICE, HRQoL is 'A combination of a person's overall physical, mental and social wellbeing; not merely the absence of disease'. For the model, QALYs were assigned to each health state. This represents the HRQoL of a patient at that health state over a period of 1 year. QALYs were estimated based on the area under the curve (AUC) of health utility estimates over the 1-year cycle length spent in each health state. Health utility values were estimated using patient-reported EQ-5D responses and applying general population preferences (i.e. a social value set or 'tariffs').¹⁹³ If study participants did not complete the EQ-5D, then responses from alternative questionnaires were mapped to it. Lower EQ-5D health utility estimates indicate worse QoL, with scores ranging between –0.594 and '1' (perfect health), with an anchor at '0' representing death. The discount rate for QALYs was 3.5% per year in accordance with NICE guidelines.¹⁹⁴

In both knee and hip models, health utility values for the primary knee (hip) replacement, post-replacement, revision and post-revision states were estimated using EQ-5D mean index from the NHS PROMs data set. Because PROMs data only cover one pre- and one postoperative measure, the utility associated with primary knee (hip) replacement and revision states were calculated by assuming patients reach the postoperative EQ-5D value at 6 months (half cycle) from the primary knee (hip) replacement/revision and maintain that health utility for the following 6 months and throughout the post-primary and post-revision health states.

Knee

The studies reporting health utility estimates for patients with early knee OA offered limited certainty, mainly because of the high risk of bias or the imprecision in the reported estimates. The evidence from the Kirkley *et al.* RCT study suggested there may be little difference between arthroscopic debridement (including some meniscal excision) and physiotherapy at 2 years' follow-up (although there are wide CIs).⁸⁸ Evidence from one comparative NRS suggested a possible benefit of surgery in terms of combined pain and function at 2 years, but this may not persist to 5 years.⁹⁴ For our model, we derived utility values for the intervention and post-intervention states from the Kirkley *et al.* study.⁸⁸ They compared arthroscopic debridement and physical and medical therapy versus physical and medical therapy with primary outcome being the total WOMAC score (range, 0–2400; higher scores indicate more severe symptoms) at 2 years of follow-up. For the mapping of WOMAC scores to EQ-5D, we used the algorithm by Barton *et al.* retrieved from the HERC curated database of mapping studies.^{195,196} Total WOMAC scores were converted to 0–96 scale range and then Model C by Barton *et al.* was applied to derive baseline, 3, 6, 12, 18 and 24 months EQ-5D estimation for patients in the control and surgery groups. SEs were derived based on the reported SDs and group sizes at each time point. Mapped EQ-5D health utility estimates and utility inputs are reported in [Figures 11 and 12](#) and [Table 22](#).

Values were then fitted to match their baseline value by calculating the average of the baseline estimates for the two groups. Absolute change in values and the percentage of potential change (PoPC) achieved was then calculated. PoPC provides a measure of change relative to the maximum that could be achieved and used the observed values for each group to rescale follow-up levels for each based on the new common baseline value.¹⁹⁷ QALYs were derived by calculating the corresponding AUC. The level of QALYs associated with the intervention state was derived by measuring the AUC to the EQ-5D health utility estimates from baseline to 12 months. QALYs associated with 'postintervention' states were obtained with the same method applied from 12 to 24 months. This AUC method assumes a linear relationship between values at different time points; that is, if the EQ-5D score changes between baseline and 3 months, a straight line would be drawn between the two points (see [Figure 11](#)).

Hip

For the comparison of non-regenerative surgical procedure versus non-operative management, three RCTs and one NRS study were retrieved from the evidence synthesis of published literature.^{103–105,107} The RCT by Mansell *et al.* reported combined pain and function outcomes, and displayed high levels of heterogeneity which may have resulted from 70% of participants switching from the conservative to the surgical intervention assessed by the study. This study was discarded and not used in the model. Two RCTs conducted in the United Kingdom reported QoL at short-term, suggesting no clinically important difference between the non-regenerative surgical and conservative treatments.^{103,105} The study by Palmer *et al.* had a smaller sample size (648 vs. 222 participants) and shorter follow-up time (eight vs. 12 months) and therefore we decided to populate the model based on the values reported in the study by Griffith *et al.*^{103,105} We tested sensitivity of results to utility values reported by Palmer *et al.* in a deterministic sensitivity analysis (DSA) ([Table 23](#)).

Baseline EQ-5D values for the arthroscopic debridement and non-operative groups were fitted to match their baseline origins and corresponding AUCs were calculated. Utility estimates associated with the post-intervention states were assumed equal to the reported EQ-5Ds scores at 12 months. SEs were derived based on the reported SDs and group sizes. Mapped hip EQ-5D health utility estimates and utility inputs are reported in [Figures 13 and 14](#) and [Table 24](#), respectively.

Resource use and costs

This evaluation was undertaken from a healthcare system perspective. As such, only costs incurred by the healthcare system were included. These costs relate to hospital (surgical/non-surgical) procedures undertaken and primary care utilisation by patients. The discount rate for costs is 3.5% per year in accordance with NICE guidelines.¹⁹⁴

Costs inputs were derived from the national reference cost (2019–20) comprising aggregated costs (the average unit cost of providing defined services to NHS patients in England) and patient-level costs (a cost based on the specific interactions a patient has, and the events related to their healthcare activity, such as drugs, nursing and theatre costs). In the absence of specific unit cost references, costs were derived from publications highlighted in the systematic review of economic studies. All unit costs are presented in Great British pounds (£) in 2019–20 prices. Where unit costs

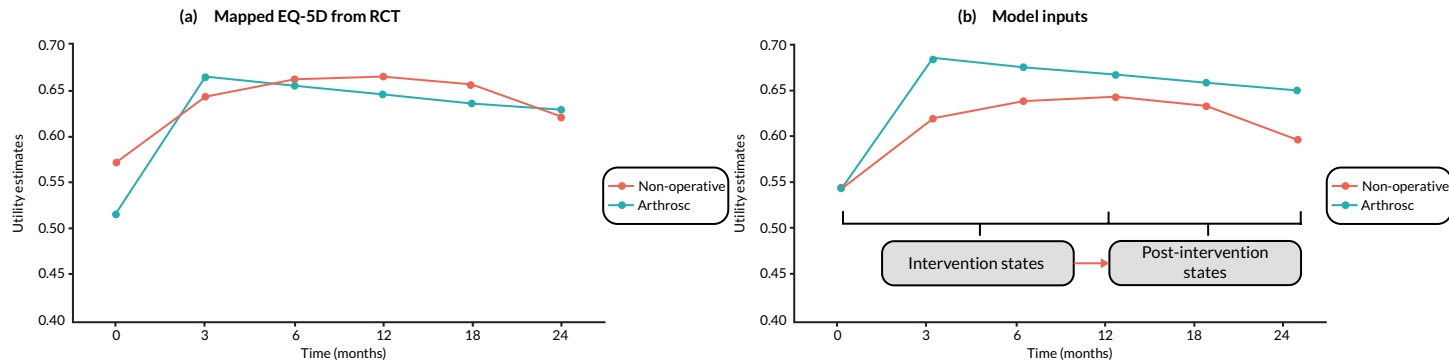


FIGURE 11 (a) Mapped EQ-5D from RCT study⁸⁸ and (b) EQ-5Ds after matching baseline values – knee model.

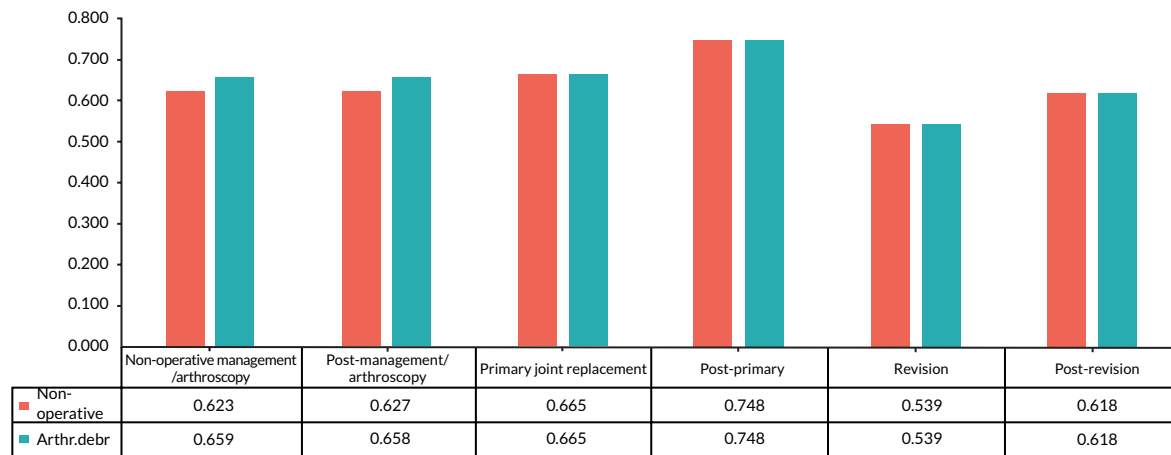


FIGURE 12 A diagram showing the utility inputs for the knee model.

TABLE 22 Utility inputs – knee model

	Health state	QALYs	SE	Source
Non-surgical arm	Non-operative management	0.623	0.045	Kirkley (2008) (baseline, 3, 6, 12 months)
	Post management	0.627	0.048	Kirkley (2008) RCT, 12, 18, 24 months
Surgical arm	Arthroscopic debridement	0.659	0.044	Kirkley (2008) RCT, linear uptake (baseline, 3, 6, 12 months)
	Post arthroscopy	0.658	0.051	Kirkley (2008) RCT, 12,18,24mo
Surgical and Non- surgical	Primary joint replacement	0.665	0.0018	PROMS (2018–9) pre- and post-operative
	Post replacement	0.748	0.0016	PROMS (2018–9) post-operative
	Revision	0.539	0.0108	PROMS (2018–9) pre- and post-operative
	Post revision	0.618	0.0105	PROMS (2018–9) post-operative

TABLE 23 Cost-effectiveness results for DSA model

Group	Total mean QALYs	Total mean cost (£)	Incremental cost (95% CI) (£)	Incremental QALYs (95% CI)	ICER (£) (cost per QALY gained)
Non-operative management	3.17	682			
Arthroscopic debridement	3.36	4306	3624.4 (3606.5 to 3645.4)	0.1894 (0.18 to 0.2)	19,134.78

were not presented in Great British pounds (£), they were inflated using the Hospital and Community Health Service (HCHS) and NHS cost Inflation Index (NHSCII) pay and prices index.¹⁹⁸

Knee

The cost estimates for the knee model are presented in [Table 25](#) and [Figure 15](#). The cost of the non-operative management was obtained from Mistry *et al.* reporting the cost of the non-surgical care package which includes paracetamol (twice a day per year) and ibuprofen (once day per year) and eight physiotherapy hospital sessions (30 minutes each).¹⁹⁹

The cost of the arthroscopic debridement (arthroscopy and cell harvest) procedure was derived from Mistry *et al.*⁴⁴ Moreover, based on the opinion of clinical experts, we assumed patients undergo one consultant-led outpatient first attendance visit and one follow-up visit (HRG:WF01A) in the same year of the arthroscopy. Costs of the corresponding Healthcare Resource Groups (HRGs) (WF01A, WF01B) were derived from national reference costs (2019–20). As per clinical expert opinion, the cost associated with the post-intervention states (post management and post arthroscopy) was assumed to be zero.

The cost of a first knee replacement was obtained from the national reference costs (HRG code: HN22E, cc score 0–1). The HRG code and complexity score were derived from Mistry *et al.*¹⁹⁹ For this health state we also considered the cost of one follow-up visit attendance (HRG: WF01A) as suggested based on consultation with clinical experts. We then assumed that in the year(s) after knee replacement patients undergo non-specialist rehabilitation services (level 3). The corresponding cost of rehabilitation associated with the post-replacement state was obtained from the national reference costs 2019–20. Patients receiving a revision (second knee replacement) have associated costs due to the operation and one consultant-led outpatient follow-up visit (HRG: WF01A).¹⁹⁹ The cost of the post-revision state was assumed to be equal to that of the post-replacement state, with patients undergoing non-specialist rehabilitation services (level 3). All costs derived from literature sources were inflated to 2019–20 prices using the HCHS/NHSCII pay and prices index.¹⁹⁸

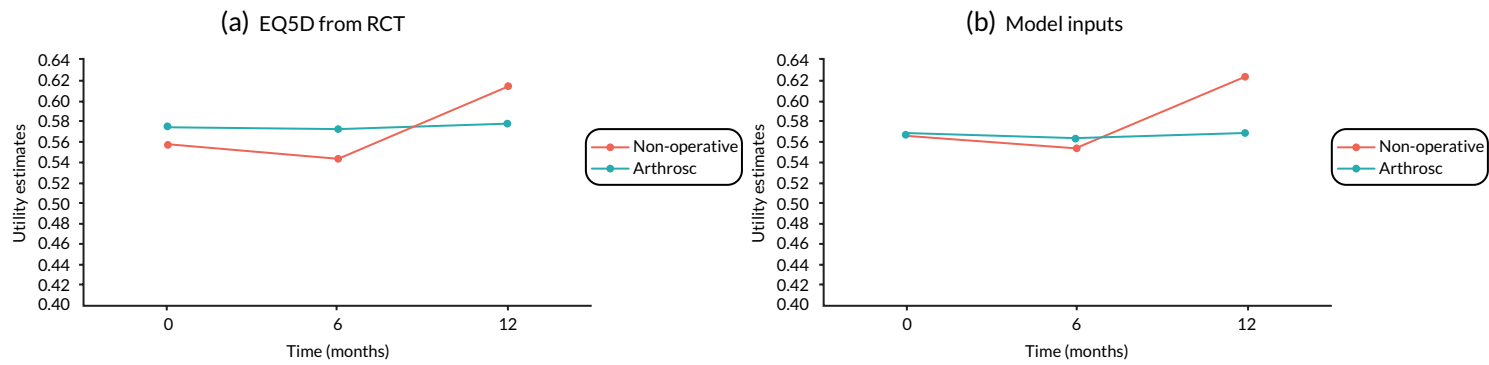


FIGURE 13 Mapped EQ-5D and utility inputs – hip model.

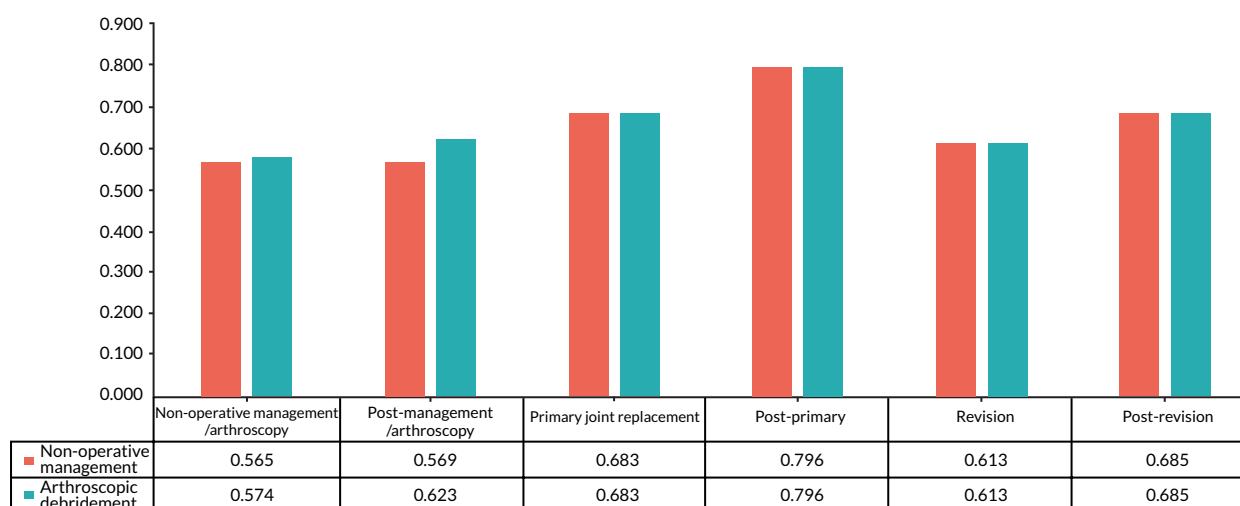


FIGURE 14 A diagram showing the utility inputs for the hip model.

TABLE 24 Utility estimates – hip model

	Health state	QALYs	SE	Source
Non-surgical arm	Non-operative management	0.565	0.018	Griffin <i>et al.</i> (baseline, 6, 12 months)
	Post management	0.569	0.018	Griffin <i>et al.</i> (12 months)
Surgical arm	Arthroscopic debridement	0.574	0.019	Griffin <i>et al.</i> (baseline, 6, 12 months)
	Post arthroscopy	0.623	0.019	Griffin <i>et al.</i> (12 months)
Surgical and Non-surgical	Primary joint replacement	0.683	0.002	PROMS (2018–9) – pre- and post-operative
	Post replacement	0.796	0.002	PROMS (2018–9) – post-operative
	Revision	0.613	0.008	PROMS (2018–9) – pre- and post-operative
	Post revision	0.685	0.008	PROMS (2018–9) – post-operative

TABLE 25 Cost inputs – knee model

	Health state	Annual cost	Source	Assumptions
Non-surgical arm	Non-operative management	£180	Mistry <i>et al.</i> (2019)	
	Post management			No interventions are provided in the post states hence no costs
Surgical arm	Arthroscopic debridement	£1085	Mistry <i>et al.</i> (2017 and 2019), reference costs 2019–20	Include one first attendance and one follow-up visit
	Post arthroscopy			No interventions are provided in the post states hence no costs
Surgical and non-surgical	Primary joint replacement	£6323	Mistry <i>et al.</i> (2019), reference costs 2019–20	Include one outpatient follow-up visits
	Post replacement	£362	Reference costs 2019–20 – non-specialist rehabilitation services level 3	
	Revision	£13,676	Mistry <i>et al.</i> (2019), reference costs 2019–20	Include one outpatient follow-up visits
	Post revision	£362		Assumed equal to post replacement



FIGURE 15 A figure showing the cost inputs for the knee model.

Hip

The cost estimates for the hip model are presented in [Table 26](#) and [Figure 16](#). The cost of the non-operative management was derived from the same RCT study we used to inform utility values in the hip model.¹⁰³ A cost of £167 was applied, corresponding to six physiotherapy sessions (average 30 minutes) in a year. No costs were associated with the year following that of the non-operative intervention (post-management state).

Patients in the arthroscopy group had an associated cost of the procedure equal to £3826, corresponding to 35% staff time, 28% surgical devices and anaesthetic drugs, 19% theatre-running costs, and 18% bed-day costs.¹⁰³ Based on consultations with clinical experts, the cost of the post-arthroscopy state included six physiotherapy sessions (same as non-operative), three follow-up visits (HRG: WF01A), and one postoperative MRI (HRG: DIM004), costs for all of which were derived from the national reference costs 2019–20.

The cost of the primary hip replacement operation was derived from the national reference cost table (HRG: HT12E, cc score 0–1). We assumed the same cc score as in the knee model and the corresponding HRG code but for the hip. One consultant-led outpatient follow-up consultation was assumed to be had by patients reaching the post-primary health state, as suggested by the clinical experts. The cost of THR revision was derived from Burn *et al.*, which was estimated based on routinely collected data in England.²⁰⁰ The cost of the post-revision state was assumed as that of one consultant-led follow-up consultation of the same type as the one provided to patients after primary replacement (HRG: WF01A). All costs were inflated to 2019–20 prices using the HCHS/NHSCII pay and prices index.¹⁹⁸

Complications

One NRS in the clinical evidence review for the knee reported no AEs in either arm up to 5 years.⁹⁴ In the case of hip interventions, very low certainty evidence suggests that in patients with a history of hip arthroscopy prior to THR, there may be a small increase in the risk of adverse effects post-THR surgery, but the evidence is highly uncertain. Therefore, AEs were considered not to be different between the comparators and hence were not included in the cost-effectiveness models.

Mortality

Probabilities of transition to death for both knee and hip models are presented in [Table 27](#). Age-specific mortality rates were based on the UK general population lifetime tables (2018–20) from the Office of National Statistics.¹⁹⁰ As the cohort ages, mortality rates generally increase throughout the time horizon in the model; therefore yearly probabilities of mortality over 5 years were derived. The starting age of the cohort in the knee model was assumed to be 59 years based on the study by Kirkley *et al.*, which was also used to inform the utility associated with the intervention states.⁸⁸ Similarly, based on Griffin *et al.* patients in the hip model were assumed to be aged 35 years when entering the 5-year simulation.¹⁰³ Based on data from the Hospital Episode Statistics database 2000–1 to 2017–8 in England, Morgan *et al.* report the prevalence of OA diagnoses in men and women.²⁰¹ To reflect the fact that the distribution was higher among

TABLE 26 Cost inputs – hip model

	Health state	Annual cost	Components	Source
Non-surgical arm	Non-operative management	£167		Griffin <i>et al.</i>
	Post management			
Surgical arm	Arthroscopic debridement	£3286		Griffin <i>et al.</i>
	Post arthroscopy	£712	£164	Six physiotherapy sessions (Griffin <i>et al.</i>)
			£375	Three follow-up visits: WF01A – reference costs 2019–20
			£173	Post-op MRI – reference costs 2019–20 (DIMOD4)
Surgical and non-surgical	Primary joint replacement	£7.545		Reference costs 2019–20 – HT12E, total unit cost
	Post replacement	£125		One consultant-led outpatient follow-up attendance (HRG code: IWF01A) – reference costs 2019–20
	Revision	£7481 (SE:615,5)		Burn <i>et al.</i>
	Post revision	£125		One consultant-led outpatient follow-up attendance (HRG code: IWF01A) – reference costs 2019–20

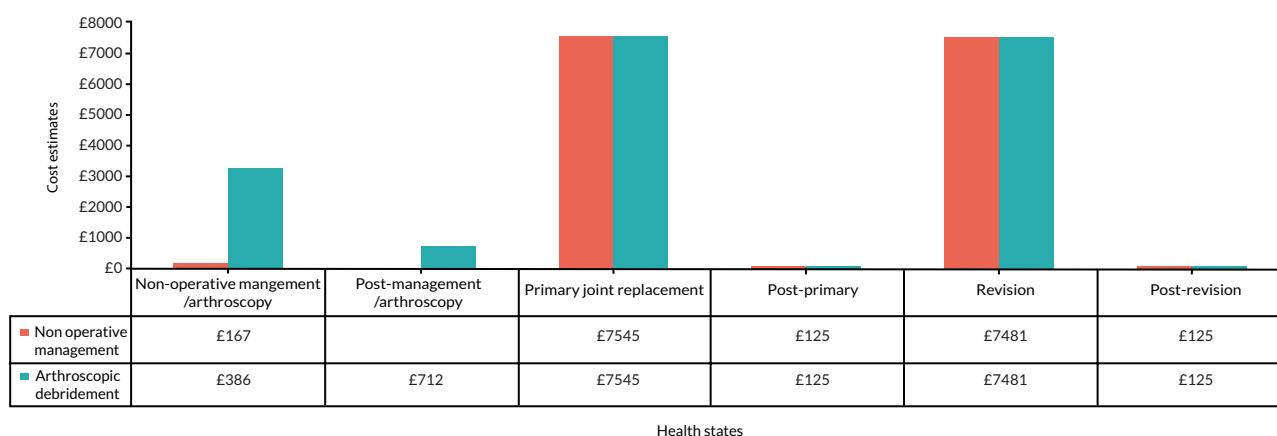


FIGURE 16 A figure showing the cost inputs for the hip model.

TABLE 27 Probabilities of transition to death for both knee and hip models

Knee model			Hip model		
Age	Cycle	Weighted rate ^a	Age	Cycle	Weighted rate
59	1	0.005796	35	1	0.000790
60	2	0.006393	36	2	0.000864
61	3	0.006926	37	3	0.000991
62	4	0.007806	38	4	0.000997
63	5	0.008424	39	5	0.001098

^a Mortality rate between age x and $(x + 1)$, that is the probability that a person aged x will die before reaching age $(x + 1)$, [10].

women than in men, we weighted age-related mortality rates by the corresponding proportion of men and women with OA of the knee (43% in men and 57% in women) and the hip (39 in men 61% in women).

Sensitivity and scenario analyses

The effect of parameter uncertainty was assessed using probabilistic sensitivity analysis (PSA), with input parameters assigned values from corresponding probability distributions and 1000 Monte Carlo simulations conducted for each subgroup. Probability distributions were based on the nature of the parameter and type of data, with gamma distributions used for costs, and beta distributions used for QoL and transition probabilities.¹⁸⁹ As the values for all the costs (except for revision THR) and transition probabilities used in the model were means or weighted averages, in order to calculate the alpha and beta values of the distributions, we assumed SEs to be 0.1 of the mean value, as has been done in previous studies.^{202,203} SEs for utility estimates were derived from the corresponding literature sources.

The sets of estimated costs and QALYs from each of 1000 Monte Carlo simulations are presented, alongside the expected results, on a cost-effectiveness plane. The simulations were also used to construct cost-effectiveness acceptability curves (CEACs) to indicate the probability of a procedure being cost-effective using a willingness-to-pay (WTP) threshold from £0 to £50,000 per QALY gained.

Knee

Distributions associated with transition probabilities, QALYs, and costs employed in the PSA for the knee model are illustrated in [Tables 28–30](#), respectively.

Scenario analysis: inputs

The literature suggested that there is an improvement in the QoL of people receiving both non-operative management and debridement. However, little difference was found between debridement and physiotherapy at short-term follow-up. The health utility of the patients who receive arthroscopy and whether the effect of arthroscopy was sustained over time are key drivers to determine whether arthroscopic debridement is the cost-effective alternative.

Given the low certainty of the evidence informing the base-case model, different scenarios on the health utility of patients who had arthroscopic debridement were tested in a scenario analysis. Specifically, we evaluated the impact of an improvement of 10% in the health utility provided by the arthroscopic debridement procedure compared to the base case. Five scenarios were built to explore the impact on results of patients sustaining the hypothetical 10% increase over 3, 6, 12, 18 and 24 months after the initial procedure.

A graphic illustration of the scenarios is shown in [Figure 17](#). The orange and blue curves represent the health utility estimate of patients in the non-operative and arthroscopy groups, respectively, as reported by the RCT source and as used in the base-case analysis. The red curves represent the hypothetical health utility progression of each of the alternative scenarios under the assumption of the 10% improvement at 3 months after arthroscopic debridement being sustained over the different periods of time.

Hip

In this section, we report the probability distributions employed in the PSA for transition probabilities, QALYs and costs ([Tables 31–33](#)) in the hip model.

Deterministic sensitivity analysis: inputs

Deterministic sensitivity analysis is a method used to investigate the sensitivity of the cost-effectiveness results to variations in a specific input parameter. We used DSA to examine the extent by which a change in the QALYs associated to the intervention and post-interventions states affects the model's cost-effectiveness output ([Figure 18](#)). This was tested by manually changing the input parameters derived from the RCT study by Griffin *et al.* to those reported by Palmer *et al.* ([Table 34](#)).^{103,105}

TABLE 28 Probability distributions for transitions – knee model

Transition from/to health states	Distribution	Mean	SE	Alpha	Beta
From non-operative management to post management					
From post management to joint replacement_cycle 1	Beta	0.021	0.002	100.0	0.000214
From post management to joint replacement_cycle 2	Beta	0.026	0.003	100.0	0.000260
From post management to joint replacement_cycle 3	Beta	0.031	0.003	100.0	0.000315
From post management to joint replacement_cycle 4	Beta	0.038	0.004	100.0	0.000381
From post management to joint replacement_cycle 5	Beta	0.046	0.005	100.0	0.000461
From arthroscopy to post arthroscopy					
From post arthroscopy to joint replacement_cycle 1	Beta	0.017	0.002	100.0	0.000167
From post arthroscopy to joint replacement_cycle 2	Beta	0.020	0.002	100.0	0.000203
From post arthroscopy to joint replacement_cycle 3	Beta	0.025	0.002	100.0	0.000246
From post arthroscopy to joint replacement_cycle 4	Beta	0.030	0.003	100.0	0.000298
From post arthroscopy to joint replacement_cycle 5	Beta	0.036	0.004	100.0	0.000361
From joint replacement to post replacement					
From post replace to revision	Beta	0.009	0.001	100.0	0.000090
From revision to post revision					
From all to death_cycle 1	Beta	0.006	0.0006	100.0	0.000058
From all to death_cycle 2	Beta	0.006	0.0006	100.0	0.000064
From all to death_cycle 3	Beta	0.007	0.0007	100.0	0.000069
From all to death_cycle 4	Beta	0.008	0.0008	100.0	0.000078
From all to death_cycle 5	Beta	0.008	0.0008	100.0	0.000084

TABLE 29 Probability distributions for QALYs – knee model

Health state	Distribution	Mean	SE	Alpha	Beta
Non-operative management	Beta	0.62	0.04	73.038	44.134
Post management	Beta	0.63	0.05	62.776	37.352
Arthroscopic debridement	Beta	0.66	0.04	75.487	38.987
Post arthroscopy	Beta	0.66	0.05	57.249	29.736
Primary joint replacement	Beta	0.67	0.00	47,046.045	23,699.887
Post replacement	Beta	0.75	0.00	52,785.991	17,765.501
Revision	Beta	0.54	0.01	1151.459	985.396
Post revision	Beta	0.62	0.01	1316.422	813.844

Results

The main outcome of the models is cost per QALY gained (represented by the ICER) for each relevant intervention/comparator pair. The ICER is calculated from outputs of the model based on the equation below:

$$ICER = \text{incremental costs} / \text{incremental QALYs}$$

TABLE 30 Probability distributions for costs – knee model

Health state	Distribution	Mean	SE	Alpha	Beta
Non-operative management	Gamma	£180	£18	100	1.80
<i>Post management</i>					
Arthroscopic debridement	Gamma	£1085	£109	100	10.85
<i>Post arthroscopy</i>					
Primary joint replacement	Gamma	£6323	£632	100	63.23
Post replacement	Gamma	£362	£36	100	3.62
Revision	Gamma	£13,676	£1,368	100	136.76
Post revision	Gamma	£362	£36	100	3.62

For healthcare interventions to be considered cost-effective (i.e. providing sufficient value for money to merit their provision) in England, they should have an incremental cost per additional QALY of less than £20,000–30,000. We now report the results for the knee and the hip models.

Knee

One thousand patients entered the model, with a starting age of 59 years. For the treatment of early OA, these patients could receive either an arthroscopic debridement or non-operative management. If symptoms were to persist, some of these patients would receive a primary knee replacement and, subsequently, potentially, a revision. [Table 33](#) shows the cost-effectiveness results over the 5-years' time horizon. Deterministic results show that arthroscopic debridement would be expected to cost £742 more than non-operative management and would generate 0.13 more QALYs than non-operative management. The cost per QALY gained for arthroscopic debridement compared with non-operative management of the knee was estimated to be £5,649.

One of the key cost drivers was the cost of the arthroscopic procedure (£1085 vs. £180 of the non-operative management), but there is also a (small) gain of QALYs from this and cost savings to the healthcare system later due to fewer people needing a primary replacement and consequently a revision.

Despite the deterministic value of the ICER being well below the WTP threshold of £20,000 per QALY, the PSA shows that in 35% of the simulations reflecting the impact of parameter uncertainty on cost-effectiveness results arthroscopic intervention would lead to lower QALY gain compared to non-operative management. The uncertainty around the cost-effectiveness results is illustrated in [Figure 19](#), showing the probabilistic distribution of the Monte Carlo simulations on the cost-effectiveness plane.

[Figure 19](#) shows the CEAC for the knee analysis. If the decision-maker were WTP £20,000 or £30,000 per QALY gain, the probability that arthroscopic debridement would be cost-effective was found to be between 62% and 63.5%. For a WTP of £50,000, the probability of being cost-effective would be 64.5%.

Scenario analysis: results

The base-case analysis revealed great uncertainty around the incremental utility provided by the arthroscopic intervention compared to the non-operative management. The scenario analysis had the objective to test the case of a greater and more durable effect over time for the arthroscopic intervention. [Figure 20](#) shows the CEACs derived by running the cost-effectiveness model under the different scenario assumptions described above. If arthroscopy provided a 10% higher utility to patients with early OA than the base case, with its effect lasting for 12 months, the probability of this procedure being cost-effective would be approximately 73% under a WTP of £20,000 per QALY gained. If the improvement were to be sustained over 18 months, the probability of cost-effectiveness would reach 90%.

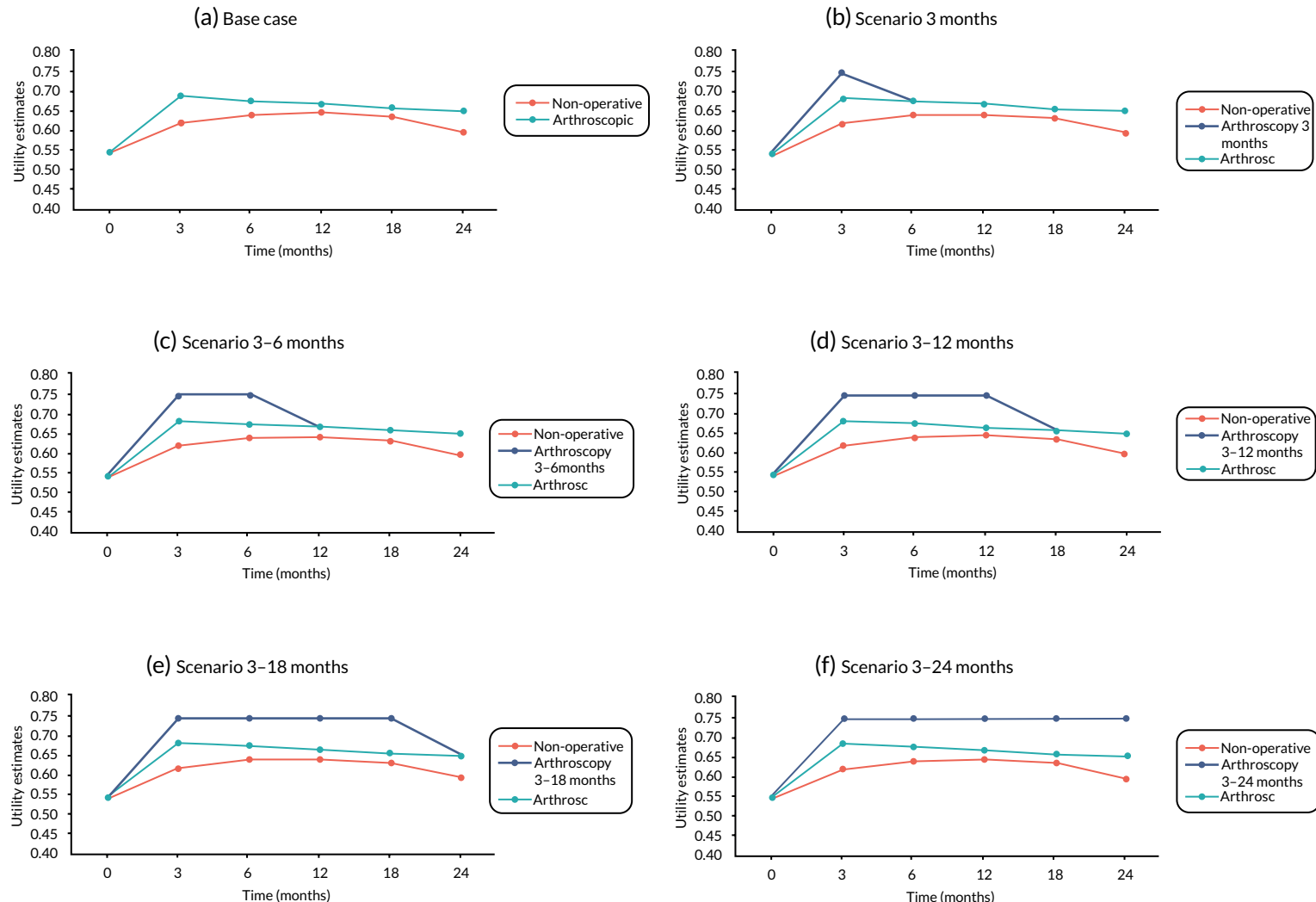


FIGURE 17 Utility estimates for intervention/post-intervention states under each scenario – knee model.

TABLE 31 Probability distributions for transitions – hip model

Transition from/to health states	Distribution	Mean	SE	Alpha	Beta
<i>From non-operative management to post management</i>					
From post management to joint replacement	Beta	0.0197	0.002	100	0.00020
<i>From arthroscopy to post arthroscopy</i>					
From post arthroscopy to joint replacement	Beta	0.022	0.002	100	0.00022
<i>From joint replacement to post replacement</i>					
From post replacement to revision (non-operative)	Beta	0.010	0.001	100	0.00010
From post replacement to revision (arthroscopy)	Beta	0.017	0.002	100	0.00017
<i>From revision to post revision</i>					
From all to death_cycle 1	Beta	0.0008	0.0001	100	0.00001
From all to death_cycle 2	Beta	0.0009	0.0001	100	0.00001
From all to death_cycle 3	Beta	0.0010	0.0001	100	0.00001
From all to death_cycle 4	Beta	0.0010	0.0001	100	0.00001
From all to death_cycle 5	Beta	0.0011	0.0001	100	0.00001

TABLE 32 Probability distributions for QALYs – hip model

Health state	Distribution	Mean	SE	Alpha	Beta
Non-operative management	Beta	0.57	0.02	413.695	317.890
Post management	Beta	0.57	0.02	427.996	324.796
Arthroscopic debridement	Beta	0.57	0.02	373.413	276.760
Post arthroscopy	Beta	0.62	0.02	399.778	241.656
Primary joint replacement	Beta	0.68	0.00	64,444.327	29,969.358
Post replacement	Beta	0.80	0.00	47,656.592	12,212.822
Revision	Beta	0.61	0.01	2569.101	1620.750
Post revision	Beta	0.69	0.01	2321.303	1066.705

TABLE 33 Probability distributions for costs – hip model

Variable	Distribution	Mean	SE	Alpha	Beta
Non-operative management	Gamma	£167	£17	100	1.67
<i>Post management</i>					
Arthroscopic debridement	Gamma	£3286	£329	100	32.86
Post arthroscopy	Gamma	£712	£71	100	7.12
Primary joint replacement	Gamma	£7545	£755	100	75.45
Post replacement	Gamma	£125	£13	100	1.25
Revision	Gamma	£7481	£748	100	74.81
Post revision	Gamma	£125	£13	100	1.25

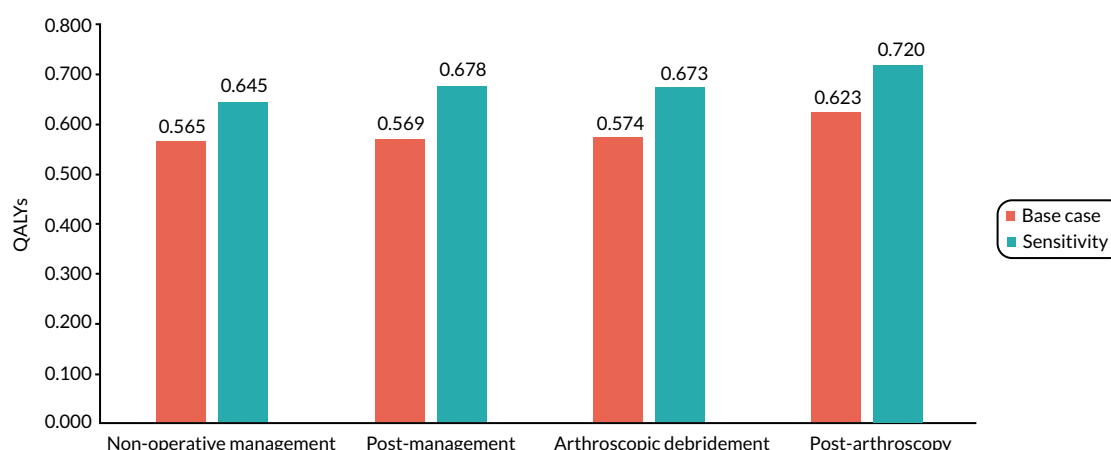


FIGURE 18 Figure showing QALY inputs for DSA for the hip model

TABLE 34 Quality-adjusted life-year inputs for DSA – hip model

	Base case (Griffin <i>et al.</i>)		Sensitivity (Palmer <i>et al.</i>)	
	QALYs	SE	QALYs	SE
Non-operative management	0.565	0.018	0.645	0.027
Post management	0.569	0.018	0.678	0.026
Arthroscopic debridement	0.574	0.019	0.673	0.027
Post arthroscopy	0.623	0.019	0.720	0.027

Hip

Table 35 shows the cost-effectiveness results for the early OA hip model. Arthroscopic debridement costed £3624 more than non-operative management and generated 0.25 more QALYs. The cost per QALY gained (ICER) was just under £15,000, suggesting that arthroscopic debridement is a cost-effective intervention compared to non-operative management.

Figure 21 shows the distribution of the probabilistic simulations on the cost-effectiveness plane. The majority of Montecarlo simulations (98%) fell in the NE quadrant indicating an expected favourable effectiveness in terms of HRQoL for those receiving arthroscopic debridement. The probability of the surgical intervention being cost-effective, under a WTP of £20,000 (£30,000) per QALY gained, was 70% (84%).

Deterministic sensitivity analysis: results

We ran a DSA on the health utility estimates underpinning the QALYs associated with the main intervention and their immediately subsequent model health states based on a different literature source (Palmer *et al.*) than the one considered for the base case (Griffin *et al.*).^{103,105} Using the alternative values, the ICER increases to £19,135 per QALY gained (**Table 36**) and so does the uncertainty around the expected effectiveness results (**Figure 22**). At £20,000 per QALY gained WTP, the probability of arthroscopic debridement being cost-effective would be, instead, 51.6%.

Discussion and conclusions

In the base-case knee analysis, the discounted deterministic results showed that arthroscopic debridement is expected to be cost-effective compared to non-operative management at an ICER of £5649 per QALY gained. However, this result comes from a clear finding of arthroscopic debridement being more costly than non-operative management but leading to a small average gain in QALYs. The mean expected gain in QALYs over the 5 years was 0.13, which is

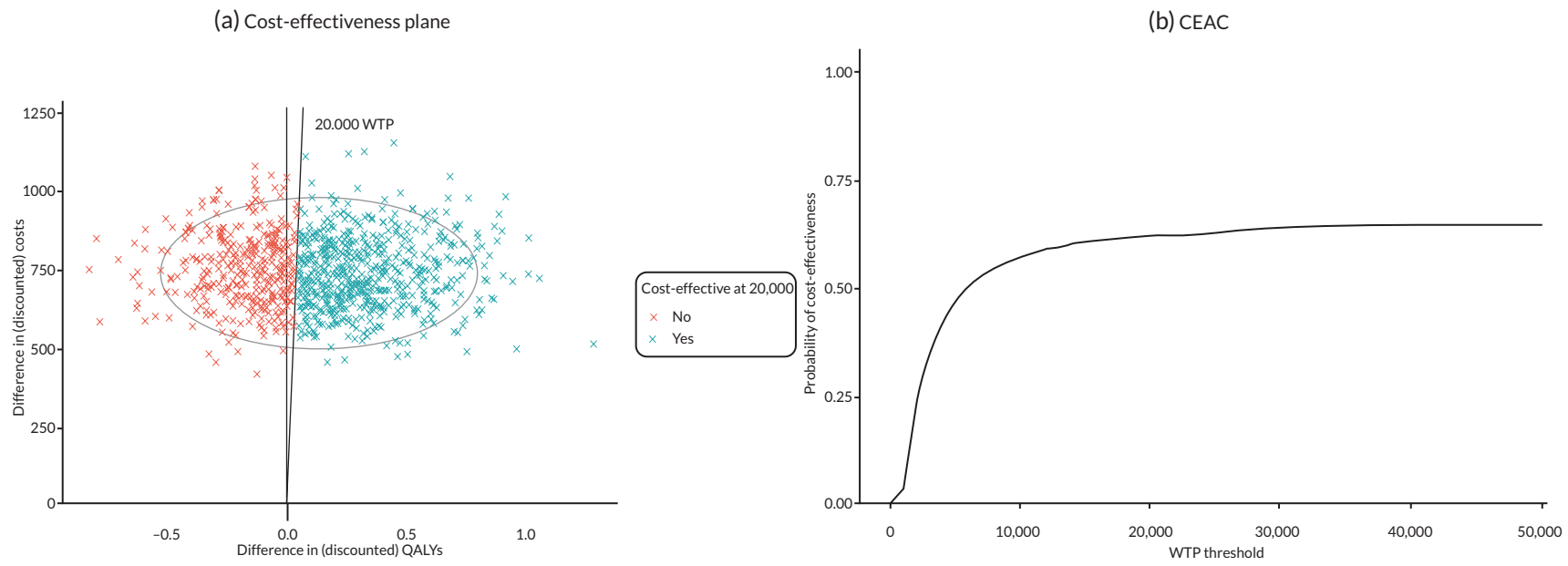


FIGURE 19 Cost-effectiveness plane (a) and CEAC (b) – knee model

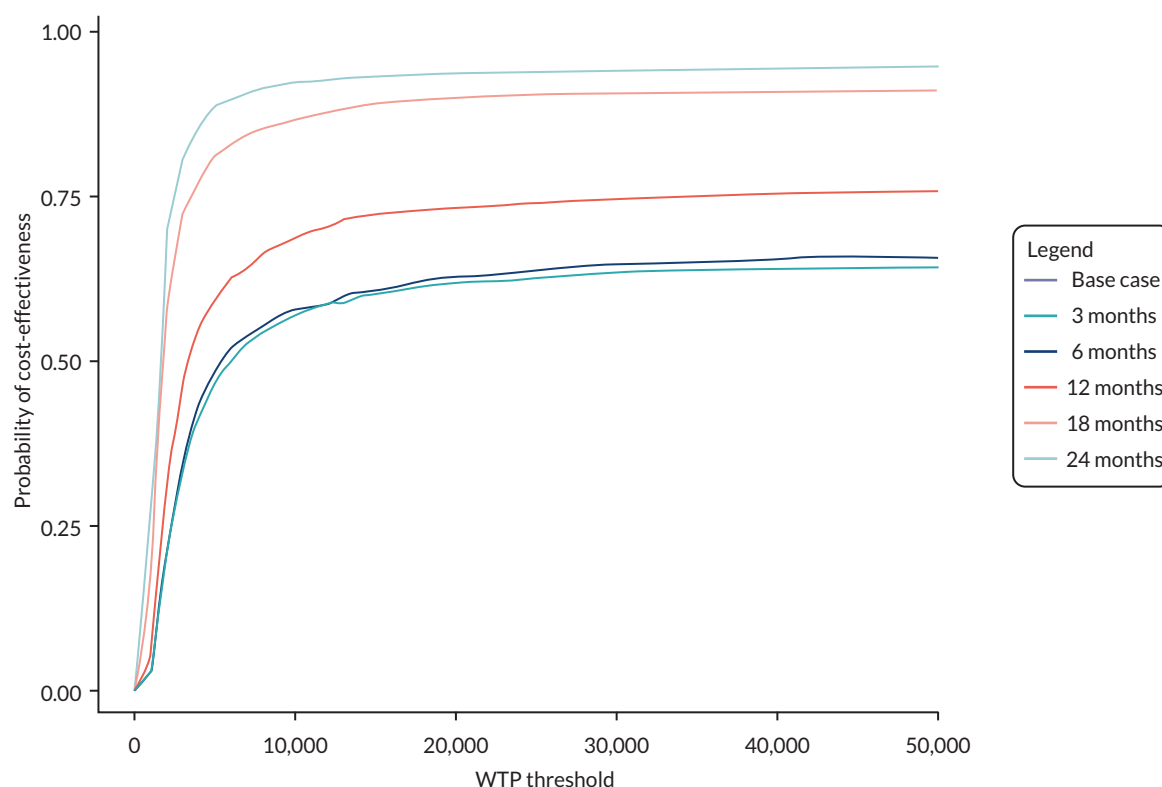


FIGURE 20 Cost-effectiveness acceptability curve base case vs. scenarios. Note: the curve corresponding to base case is not notable in the figure, but it is simply sitting underneath that of the 3-month scenario.

TABLE 35 Cost-effectiveness result for the hip model

Group	Total mean QALYs	Total mean cost (£)	Incremental cost (95% CI) (£)	Incremental QALYs (95% CI)	ICER (£) (cost per QALY gained)
Non-operative management	2.68	682			
Arthroscopic debridement	2.93	4306	3624.4 (3606.5 to 3645.4)	0.25 (0.24 to 0.26)	14,602.37

low but unsurprising considering that the expected postoperative health utility estimate was only 0.02 greater than that of post-non-operative management. This difference is nearly nine times lower than the MCID identified for patients of knee replacement in England and hence might be suggestive of a measured improvement that does not quite reach levels such that it would make a difference to patients.²⁰⁴ Furthermore, there is great uncertainty around the cost-effectiveness results, especially around the QALYs differences. Full PSA shows that the uncertainty around model parameters place the probability of cost-effectiveness of arthroscopic debridement at 62% when considering a threshold of £20,000 per QALY gain, with this increasing to only 65% at a threshold of £50,000. This suggests there is approximately a one in three chances that arthroscopic debridement is actually not cost-effective, on the basis of surgery leading to yet smaller gains or even a loss in QALYs compared to non-operative management. Based on the sensitivity of cost-effectiveness results to the relative QALY impact of arthroscopic debridement, we explored various hypothetical scenarios of surgery leading to higher health utility and this being sustained for different periods of time. If arthroscopic debridement were to achieve a 10% higher health utility after surgery than what we identified in the evidence review and used in the base-case analysis, with this extra gain being sustained for 2 years, the probability of cost-effectiveness at a threshold of £20,000 per QALY gained would reach 90%.

In the case of the hip analysis, our model found that, in the deterministic analysis, arthroscopic debridement was expected to be cost-effective compared to non-operative management with an ICER of £14,602 per QALY gained.

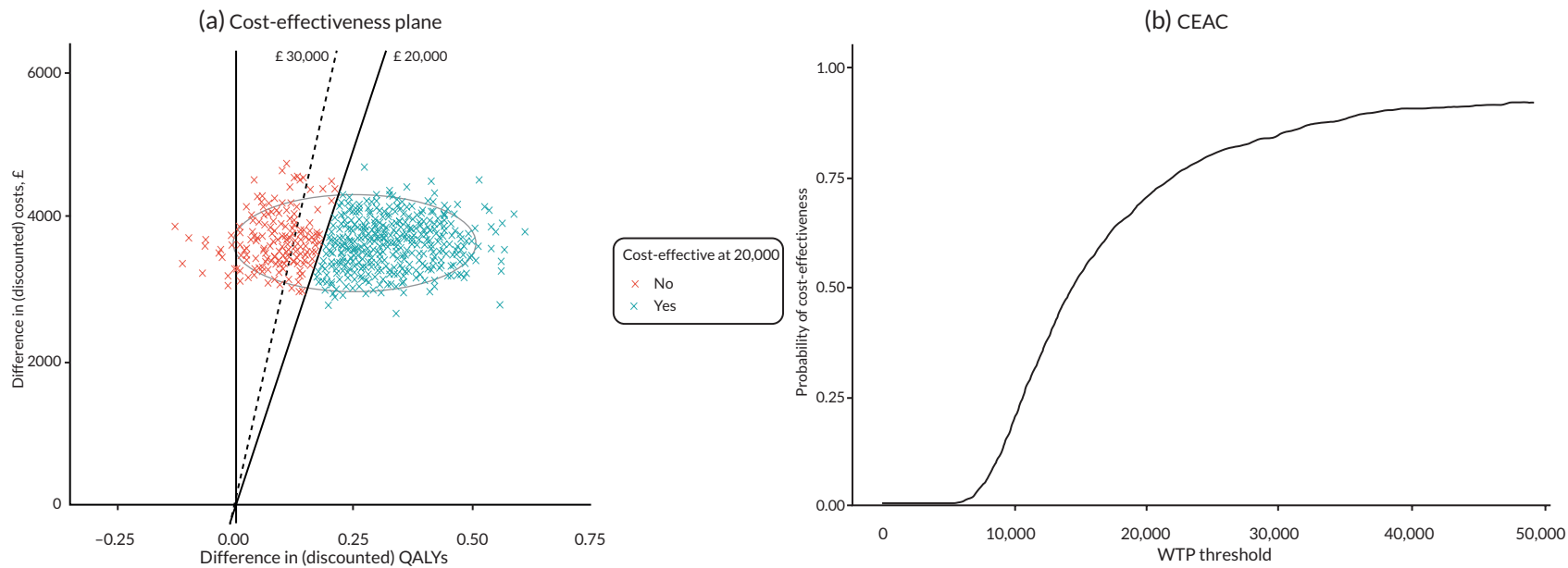


FIGURE 21 Cost-effectiveness plane (a) and CEAC (b) – hip model.

TABLE 36 Cost-effectiveness result for the knee model

Group	Total mean QALYs	Total mean cost (£)	Incremental cost (95% CI) (£)	Incremental QALYs (95% CI)	ICER (£) (cost per QALY gained)
Non-operative management	2.90	972.00			
Arthroscopic debridement	3.03	1715.00	742.14 (734.4 to 748.5)	0.13 (0.11 to 0.15)	5649.25

As expected, surgery is expected to report higher costs (£3624) but also to lead to a QALY benefit of 0.25 over the 5 years. This expected gain in QALYs was achieved by feeding the model with estimated health utility values that were 0.05 greater for surgery than for non-operative management, a difference far lower than the 0.19 identified as the minimum to be clinically important for patients.²⁰⁴ PSA showed that, considering uncertainty in all parameters used for the simulation, there is a 70% probability that arthroscopic debridement would be cost-effective compared to non-operative management. However, as there was also uncertainty not only around the parameter values but also around the evidence and mean health utility values populating the model, another SA using study findings by Palmer *et al.* as opposed to those by Griffin *et al.* was conducted.^{103,105} The main difference between these studies is that the one reported in Palmer *et al.* had a smaller sample size (222 participants, compared to 648 in Griffin *et al.*) and shorter follow-up time (8 months compared to 12).¹⁰⁵ When using the values reported by Griffin *et al.*, the expected (deterministic) ICER reached £19,134, that is nearly exactly on the more commonly used threshold of £20,000, and hence the probability of cost-effectiveness at said threshold dropped to 51.6% suggesting it would be essentially as likely that arthroscopic debridement would be cost-effective as that it would not.¹⁰³

The economic models we employed had several limitations. First, the review of existing clinical studies found scarce evidence specific to the comparisons of interest and with high risk of bias. For example, the length of the follow-up data found was too short; we addressed this by making several assumptions to associate utility estimates with post-intervention states. Also, we had to rely heavily on a few studies, such as Kirkley *et al.* for the knee and Griffin *et al.* for the hip model.^{88,103} As a result, we had to make several assumptions, each of which we checked for plausibility with clinical experts. Second, we also relied on clinical experience for information on the average resources used (e.g. number of outpatient visits) for patients in both models, which may underestimate the actual use of resources by patients in different settings. NICE's guide to the methods of technology appraisal suggests a perspective of analysis of the NHS and personal social services, but we were not able to include personal social services as no evidence was identified reporting the impact of arthroscopic debridement on personal social services vis à vis non-operative management.²⁰⁵ Other costs such as productivity losses and informal care could potentially be relevant when comparing these interventions and should be explored in future studies. We did not consider any costs or dis-utilities associated with AEs after surgery as we found this to be rarely if ever even mentioned in studies comparing arthroscopic debridement to non-operative management. Finally, given the lack of detailed reporting on costs and transition probabilities, SEs were assumed at 10% of the mean, which is a common assumption in health economic modelling practice, and the models assumed no correlations between the input distributions for the PSA given the absence of evidence for it.^{201,202}

To conclude, for both the knee and hip analyses, it appears possible that arthroscopic debridement could be cost-effective compared to non-operative management. However, the evidence around the magnitude of the gain in QoL and its uncertainty needs to be more robustly ascertained for a more conclusive cost-effectiveness finding to be produced.

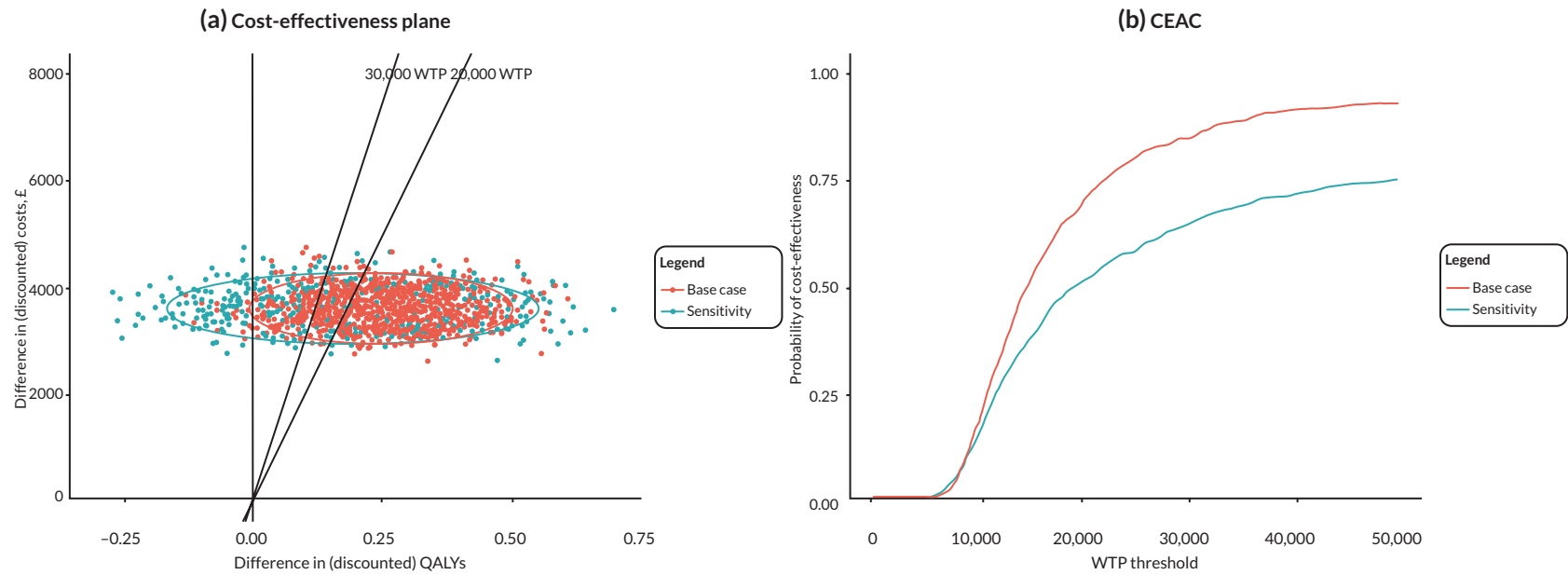


FIGURE 22 Cost-effectiveness plane (a) and CEAC (b) – DSA hip model.

Chapter 12 Discussion and research recommendations

Introduction

The aim of this study was to investigate the evidence for surgical treatments of early hip and knee OA. Many patients suffer from hip and knee OA and can be managed successfully with a non-operative approach as laid out by the NICE guidance for treating OA. In some patients a non-operative approach does not control the disabling OA symptoms. For those patients with severe disease, characterised by complete loss of articular cartilage, joint replacement provides a successful intervention with a large reduction in pain, improved mobility and enhanced QoL. However, in early stages of the disease (non bone-on-bone change), a treatment gap exists for some patients, where arthroplasty is not indicated but non-operative measures fail to control symptoms. Many surgical procedures have been introduced to address this gap but very few are widely used, and none are included in the NICE guidance for managing OA of the hip and knee due to a lack of evidence. This study plans to review these procedures to identify what evidence does exist and what evidence gaps need to be addressed by new research.

The discussion will focus initially on some general findings which are pertinent to all studies included in the work.

Definition of early osteoarthritis

For this study, we have used a well-recognised definition of early OA based on the level structural change in the joint identified by imaging methods. The definition is based on imaging methodology that grades OA change by the degree of articular cartilage damage in the joint identified by radiographs or MRI scanning. Unfortunately, the evidence synthesis did reveal that many clinical studies do not record in a detailed manner a description of the level of early OA disease they are treating. A more standardised approach should be approached if new research is undertaken so that comparisons can be made between studies.

Comparator interventions

The evidence synthesis was aimed at identifying studies that compared surgical interventions for early hip and knee OA, with a non-operative approach to managing the joint disease. We defined the non-operative comparator intervention based on NICE guidance for managing OA. Unfortunately, there were very few studies that compared a surgical intervention for early OA against established non-operative measures. Some comparative studies investigated the effect of one surgical intervention against another. As a result, in many of these studies do not provide any evidence of the effectiveness of the surgery compared with the current treatment that is routinely provided for patients. Non-operative NICE guidance-based treatment is effective for many patients, and the effect of surgical treatments must be compared with this. The interventions must be found to have a clinical benefit above non-operative measures if they are to be established as a routinely provided care. Any further research commissioned in this area must ensure that surgical procedures are compared with established and clearly defined non-operative treatments.

Levels of evidence

The systematic review component of this study was aimed at identifying a level of evidence for surgical procedures that could be meaningfully taken forward into a health economic analysis. The criteria for taking studies forward to the health economic analysis allowed inclusion of (1) RCTs of surgical treatment for early OA compared to non-operative treatment and (2) non-randomised cohort studies where a surgical intervention was compared to non-operative treatment. This methodology was developed so that the HE analysis could be driven with robust data. While some

relevant evidence was identified, in general the systematic reviews revealed that there was a major lack of studies that could provide the level of evidence required.

Types of surgical procedure

The literature reveals that there are many different recorded types of surgical intervention for early OA in both joints. They will be discussed in the four major categories of surgical procedures that we have used as a structure for this evidence synthesis: (a) non-regenerative procedures, (b) load-modifying procedures, (c) biological reconstruction procedures and (d) arthroplasty (joint replacement procedures). The discussion will now address each category for each joint, making research recommendations in each case.

Surgical procedures for treatment of early knee osteoarthritis

Non-regenerative surgical procedures for treatment of early knee osteoarthritis

There is good evidence to suggest that arthroscopic washout procedures are not effective in treating early knee OA. The most recent NICE guidance regarding this surgical procedure reports that this procedure should not be used for this indication and we found no new evidence to suggest that the conclusions reached in the treatment guidance are wrong.

There is some relatively weak evidence to suggest that arthroscopic chondral debridement could be an effective treatment. Limited comparative data from one RCT and one NRS suggest that a small clinical benefit for patients could be found using the procedure. This evidence was taken forward to the health economic analysis, which showed that there was significant uncertainty as to if this would be cost-effective. Further, well-conducted RCTs are required to establish if this surgery is an effective treatment for early knee OA.

A number of RCT studies were identified that examined the effectiveness of meniscectomy to treat patients with meniscal tears. In some of these studies, patients have early knee OA, representing an indirect treatment for early knee OA. Analysis revealed that the surgery did not appear to be an effective treatment to treat meniscal tears with or without early knee OA.

Load-modifying surgical procedures for treatment of early knee osteoarthritis

There was limited evidence to support the use of load-modifying procedures to treat early knee OA. The procedure with the most evidence was HTO with a number of studies comparing different types of surgery. At this time, there was no compelling evidence that it was more effective than non-operative treatment to treat early knee OA. We identified one RCT study that was in progress addressing this question directly, and the results of this work will be a good contribution to the limited body of evidence available at present. Further, well-conducted RCTs are required to establish if this surgery is an effective treatment for early knee OA.

Biological reconstruction surgical procedures for treatment of early knee osteoarthritis

Our systematic review revealed no direct evidence of the use of biological reconstruction to treat early knee OA compared to non-operative treatment. In one study AMIC and MF procedures were compared, without a control arm of treatment with non-operative measures. Both procedures appeared to produce some clinical improvement, but this is very weak evidence of their effect.

Our work did identify a number of surgical procedures and techniques have been developed to treat isolated cartilage defects in the knee but that have not been directed towards treating early knee OA. While isolated cartilage defects are an accepted risk factor for developing early knee OA, this type of treatment has not been adopted or investigated as an intervention where the indication is to stop progression to early OA.

There is a clear need for more research on this area, specifically producing feasibility work leading to RCT studies to investigate if the procedures can be used to treat early knee OA, with an effect that is better or at least equal to non-operative measures.

Joint replacement surgical procedures for treatment of early knee osteoarthritis

Whilst joint replacement is a very established treatment of late-stage knee OA, there are no robust studies that have directly investigated its effect when used to treat early knee OA, comparing its use to non-operative treatment. We found a single RCT study that compared HTO to UKR in this patient group. In the study, UKR produces some benefit for patients, but there is no comparison with non-operative treatment. One RCT that did compare knee arthroplasty to non-operative treatment included some patients who met our definition of early and so does offer some indirect weak evidence of a positive effect for patients. However, there is a clear need for further research on this area and well-conducted RCTs are required focused on comparing joint replacement against non-operative treatment in this patient group. However, some caution is required as we did identify prognostic studies that compare the clinical effect of knee replacement in patients with early of later stage knee OA. These studies provide some evidence that suggests the large positive effect sizes seen in patients with late-stage OA is reduced in the early stage of the disease, although this evidence is weak.

Surgical procedures for treatment of early hip osteoarthritis

Non-regenerative surgical procedures for treatment of early hip osteoarthritis

The major intervention that we identified in this category was the arthroscopic osteochondroplasty to treat FAI. This is a condition seen as the part of the early morphological changes seen in hip arthritis and is included within the scope of this work. The surgery is already widely used within the NHS. There have been three RCTs that have compared this surgery to non-operative care, with evidence produced that could be taken forward to the health economic analysis. There appears to be a clinical benefit for patients at 12 months post surgery, and following the health economic analysis, the intervention proved to be cost-effective. It appears that this intervention provides cost-effective benefit for patients supporting its use in patients where a non-operative treatment approach is not successful.

Another surgical procedure used in this group was arthroscopic labral surgery used to treat labral tears. There is no evidence that shows labral tear is directly related to the development of structural OA in the hip. However, labral tears are thought to produce symptoms and may predispose patients to an increased risk of developing hip OA at a later stage. At present, there are no studies that compare labral tear surgery with non-operative management of the condition. However, we are aware of several ongoing studies that directly address this issue.

Load-modifying surgical procedures for treatment of early hip osteoarthritis

There is no convincing evidence that shows the benefit of load-modifying procedures compared to non-operative treatment options to treat patients with early hip OA. From the additional literature reviewed in this study a number of studies identified PAO as a load-modifying procedure that is used to treat hip dysplasia, which is believed to be an early stage in the development of hip OA. However, there are no comparative studies that show its benefit when compared to non-operative measures and therefore the evidence for the clinical effectiveness of PAO is very weak. There is a real need for further research in this area leading to a well-conducted RCT of PAO against non-operative care in patients with hip dysplasia.

Biological reconstruction surgical procedures for treatment of early hip osteoarthritis

Biological reconstruction surgical procedures are used to treat patients with FAI, but there are no studies that have investigated their clinical effectiveness compared to non-operative measures. There are a series of NRS that have compared different surgical techniques used to treat FAI. Techniques reported includes MF, AMIC, membrane autologous chondrocyte implantation and hip arthroscopy BMA. Therefore, the evidence for their efficacy is weak with no comparative studies against non-operative management.

There is a real need for feasibility work leading to well-conducted RCTs of biological reconstruction techniques to treat FAI.

Joint replacement surgical procedures for treatment of early hip osteoarthritis

Hip replacement is highly successful in treating late-stage hip OA. However, this review identified no studies that have focused on hip replacement as a treatment for early hip OA (as defined in this study) that comparing its use to non-operative treatment.

As with the knee, there are a number of cohort studies that do show hip replacement is being used to treat patients with early hip OA. The data prevented suggest the outcomes for patients with early hip OA are potentially worse than for those with established late-stage OA. Therefore, as with the knee, further research is required in this area comparing hip joint replacement against non-operative treatment in patients with early hip OA group.

Patient and public involvement

We have involved patients and the public in the development of this proposal in three ways; development of the research questions, development of this research application and delivery of the study.

In planning this study, our research group initially worked with small focus groups of patients and the public to understand their thoughts around the surgical treatment of hip and knee OA. From this initial work we recognised that patients and the public raised questions that current research direction was not addressing. From work under the leadership of Professor Price (principal investigator for this application) we commissioned two James Lind Alliance (JLA) Priority Setting Partnerships (PSP) looking at surgery for treating hip and knee OA. The first dealt with joint replacement surgery and the second focused on surgical treatment of early hip and knee OA. The results of the JLA PSP looking at 'Surgical treatment of early OA' produced a top 10 research questions that were prioritised. This HTA commissioned research call has been developed specifically with reference to the JLA PSP that we set up and ran with patients driving the generation of research questions.

In developing the application for funding for this work, we continued our involvement of patients and the public. In order to maintain continuity of involvement, we recruited one of the original patient representatives from the Steering Group that oversaw the JLA PSP for 'Surgical treatment of early OA', a patient and carer with experience of early hip and knee OA. They acted as a co-applicant for funding application and were involved in production of the research plan and the detailed protocol. They had a major involvement in the production of the [Plain language summary](#).

In completing this work, our PPI representative was part of the Study Management Group (SMG) and so had a role in running and delivery of the work. The SMG had input into all aspects of the work.

Equality, diversity and inclusion

In the process of planning and conducting this work, we have been mindful to address equity, diversity and inclusion into the work. The fundamental research question is applicable to all patient populations and was developed through a JLA process where the work group contained members from a very diverse background. The process of creating the group was inclusive and allowed equitable access to patients and public. The study itself employed search strategies that would include patient populations from a diverse background. The findings of the work are applicable to all patient groups.

Limitations of research

This research has taken a long time to complete the reporting stage. As a result, a significant period of time has elapsed since the original updated searches completed in January 2021. To address this, a further search for RCTs of any surgical treatment for early hip and knee OA was completed in January 2025. Therefore, there may be new evidence relating to non-randomised comparative or non-comparative studies assessing treatments for early hip and knee OA that are not reported in this study. However, given that no new RCT is involved data, the impact of this on the summary of results is likely to be minimal, and the findings of this study will still be relevant.

Research recommendations

1. Further RCT studies are required to determine if arthroscopic chondral debridement is an effective treat for early knee OA.
2. The results of an in-progress RCT of HTO compared to non-operative treatment measures are awaited. If this shows a benefit in terms of clinical and cost-effectiveness, then further study will be required to study new technology that potentially improves the effect of the surgical procedure.
3. Feasibility studies leading to further RCTs are required to determine if ACI and other biological surgical treatments are effective treatments for early knee or hip OA when compared to non-operative treatment.
4. Feasibility studies leading to further RCTs are required to determine if PAO is a clinically effective treatment for early hip OA when compared to non-operative treatment.
5. Feasibility studies leading to further RCTs are required to determine if any of the commonly used biological surgical treatments are effective treatments for early knee or hip OA when compared to non-operative treatment.
6. Well-conducted feasibility studies are required to determine if joint replacement is a potentially effect treatment for early-stage knee or hip OA, in comparison to non-operative treatment. If a positive effect is identified, then further RCT studies will be required.

Additional information

CRediT contribution statement

Andrew Price (<https://orcid.org/0000-0002-4258-5866>): Conceptualisation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualisation, Writing – original draft, Writing – reviewing and editing.

Emma Searle: Data curation, Formal analysis, Investigation, Methodology, Project administration, Validation, Visualisation, Writing – original draft.

Maggie Westby (<https://orcid.org/0000-0003-4273-3942>): Data curation, Formal analysis, Investigation, Methodology, Project administration, Validation, Visualisation, Writing – original draft.

Gianluca Fabiano (<https://orcid.org/0000-0001-9922-4810>): Formal analysis, Investigation, Methodology, Project administration, Validation, Visualisation, Writing – original draft.

Pamela Royle (<https://orcid.org/0000-0001-5175-1005>): Data curation, Formal analysis, Investigation, Methodology, Project administration, Validation, Visualisation, Writing – original draft.

Jill Colquitt (<https://orcid.org/0000-0001-5962-2689>): Data curation, Formal analysis, Investigation, Methodology, Project administration, Validation, Visualisation, Writing – original draft.

Emma Loveman (<https://orcid.org/0000-0001-8226-2634>): Data curation, Formal analysis, Investigation, Methodology, Project administration, Validation, Visualisation, Writing – original draft.

Karen Barker (<https://orcid.org/0000-0001-9363-0383>): Conceptualisation, Funding acquisition, Methodology, Project administration, Writing – original draft, Writing – reviewing and editing.

Sion Glyn-Jones (<https://orcid.org/0000-0002-8521-7342>): Conceptualisation, Funding acquisition, Methodology, Project administration, Validation, Visualisation, Writing – original draft, Writing – reviewing and editing.

Antony Palmer (<https://orcid.org/0000-0003-4616-7482>): Conceptualisation, Funding acquisition, Methodology, Project administration, Validation, Visualisation, Writing – original draft, Writing – reviewing and editing.

Hema Mistry (<https://orcid.org/0000-0002-5023-1160>): Conceptualisation, Funding acquisition, Methodology, Project administration, Validation, Visualisation, Writing – original draft, Writing – reviewing and editing.

Toby Smith (<https://orcid.org/0000-0003-1673-2954>): Conceptualisation, Funding acquisition, Methodology, project administration, Validation, Visualisation, Writing – original draft, Writing – reviewing and editing.

Andy Metcalfe (<https://orcid.org/0000-0002-4515-8202>): Conceptualisation, Funding acquisition, Methodology, Project administration, Validation, Visualisation, Writing – original draft, Writing – reviewing and editing.

Rafael Pinedo-Villanueva (<https://orcid.org/0000-0002-4723-5128>): Conceptualisation, Funding acquisition, Data curation, Formal analysis, Investigation, Methodology, Project administration, Supervision, Validation, Writing – original draft, Writing – reviewing and editing.

Sally Hopewell (<https://orcid.org/0000-0002-6881-6984>): Conceptualisation, Funding acquisition, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualisation, Writing – original draft, Writing – reviewing and editing.

Acknowledgements

Evidence synthesis: we would like to thank Norman Waugh for his contribution to this work in its planning and execution.

Health economic analysis: we are grateful to Ian Koblbauer for his methodological advice and support in the generation of model parameter estimates for the transition to joint replacement in the knee model. His support was instrumental in estimating annual transition probabilities for the primary joint replacement health state by digitising time-to-event and creating pseudo individual-level data so that log-cumulative hazards plots could be generated and parametric models fitted and assessed for goodness-of-fit.

Data-sharing statement

All data requests should be submitted to the corresponding author for consideration. Access to anonymised data may be granted following review.

Ethics statement

This study was an evidence synthesis and health economic analysis of published data and as such did not require ethical approval.

This research was carried out following the Declaration of Helsinki ethical principles for medical research involving human participants published in 2024.

Information governance statement

This study did not handle any personal information.

Disclosure of interests

Full disclosure of interests: Completed ICMJE forms for all authors, including all related interests, are available in the toolkit on the NIHR Journals Library report publication page <https://doi.org/10.3310/TNFD7352>.

Primary conflicts of interest: Andrew Price has been a member of the HS&DR Funding Committee (Seacole).

Karen Barker: Trial Steering Committee Chair TREADON NIHR funded trial.

Hema Mistry is a member of HTA General Committee.

Toby Smith was an HS&DR Associate Board Member.

Sally Hopewell was a member of HTA Clinical Evaluation and Trials Committee.

Disclaimer

Every effort has been made to obtain the necessary permissions for reproduction, to credit original sources appropriately and to respect copyright requirements. However, despite our diligence, we acknowledge the possibility of unintentional omissions or errors, and we welcome notifications of any concerns regarding copyright or permissions.

References

1. Murray CJ, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, *et al.* Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012;**380**:2197–223.
2. Cross M, Smith E, Hoy D, Nolte S, Ackerman I, Fransen M, *et al.* The global burden of hip and knee osteoarthritis: estimates from the global burden of disease 2010 study. *Ann Rheum Dis* 2014;**73**:1323–30.
3. Murphy L, Schwartz TA, Helmick CG, Renner JB, Tudor G, Koch G, *et al.* Lifetime risk of symptomatic knee osteoarthritis. *Arthritis Rheum* 2008;**59**:1207–13.
4. Kim C, Linsenmeyer KD, Vlad SC, Guermazi A, Clancy MM, Niu J, Felson DT. Prevalence of radiographic and symptomatic hip osteoarthritis in an urban United States community: the Framingham osteoarthritis study. *Arthritis Rheumatol* 2014;**66**:3013–7.
5. Clohisy JC, Dobson MA, Robison JF, Warth LC, Zheng J, Liu SS, *et al.* Radiographic structural abnormalities associated with premature, natural hip-joint failure. *J Bone Joint Surg Am* 2011;**93**:3–9.
6. Thomas GE, Palmer AJ, Batra RN, Kiran A, Hart D, Spector T, *et al.* Subclinical deformities of the hip are significant predictors of radiographic osteoarthritis and joint replacement in women. A 20 year longitudinal cohort study. *Osteoarthritis Cartilage* 2014;**22**:1504–10.
7. Agricola R, Heijboer MP, Roze RH, Reijman M, Bierma-Zeinstra SM, Verhaar JA, *et al.* Pincer deformity does not lead to osteoarthritis of the hip whereas acetabular dysplasia does: acetabular coverage and development of osteoarthritis in a nationwide prospective cohort study (CHECK). *Osteoarthritis Cartilage* 2013;**21**:1514–21.
8. Agricola R, Waarsing JH, Arden NK, Carr AJ, Bierma-Zeinstra SM, Thomas GE, *et al.* Cam impingement of the hip—a risk factor for hip osteoarthritis. *Nat Rev Rheumatol* 2013;**9**:630–4.
9. Jones LD, Bottomley N, Harris K, Jackson W, Price AJ, Beard DJ. The clinical symptom profile of early radiographic knee arthritis: a pain and function comparison with advanced disease. *Knee Surg Sports Traumatol Arthrosc* 2016;**24**:161–8.
10. Dieppe PA, Lohmander LS. Pathogenesis and management of pain in osteoarthritis. *Lancet* 2005;**365**:965–73.
11. Glyn-Jones S, Palmer AJ, Agricola R, Price AJ, Vincent TL, Weinans H, Carr AJ. Osteoarthritis. *Lancet* 2015;**386**:376–87.
12. National Institute for Health and Care Excellence. *Osteoarthritis: Care and Management*. NICE clinical guideline CG226. London: NICE; 2022.
13. Skou ST, Roos EM, Laursen MB. A randomized, controlled trial of total knee replacement. *N Engl J Med* 2016;**374**:692.
14. Carr AJ, Robertsson O, Graves S, Price AJ, Arden NK, Judge A, Beard DJ. Knee replacement. *Lancet* 2012;**379**:1331–40.
15. Pivec R, Johnson AJ, Mears SC, Mont MA. Hip arthroplasty. *Lancet* 2012;**380**:1768–77.
16. London NJ, Miller LE, Block JE. Clinical and economic consequences of the treatment gap in knee osteoarthritis management. *Med Hypotheses* 2011;**76**:887–92.
17. Bayliss LE, Culliford D, Monk AP, Glyn-Jones S, Prieto-Alhambra D, Judge A, *et al.* The effect of patient age at intervention on risk of implant revision after total replacement of the hip or knee: a population-based cohort study. *Lancet* 2017;**389**:1424–30.
18. Dowsey MM, Nikpour M, Dieppe P, Choong PF. Associations between pre-operative radiographic changes and outcomes after total knee joint replacement for osteoarthritis. *Osteoarthritis Cartilage* 2012;**20**:1095–102.

19. Niinimäki TT, Murray DW, Partanen J, Pajala A, Leppilähti JI. Unicompartmental knee arthroplasties implanted for osteoarthritis with partial loss of joint space have high re-operation rates. *Knee* 2011;**18**:432–5.
20. Mota RE. Cost-effectiveness analysis of early versus late total hip replacement in Italy. *Value Health* 2013;**16**:267–79.
21. Moseley JB, O'Malley K, Petersen NJ, Menke TJ, Brody BA, Kuykendall DH, *et al.* A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med* 2002;**347**:81–8.
22. Sihvonen R, Paavola M, Malmivaara A, Itala A, Joukainen A, Nurmi H, *et al.*; Finnish Degenerative Meniscal Lesion Study (FIDELITY) Group. Arthroscopic partial meniscectomy versus sham surgery for a degenerative meniscal tear. *N Engl J Med* 2013;**369**:2515–24.
23. Amin NH, Hussain W, Ryan J, Morrison S, Miniaci A, Jones MH. Changes within clinical practice after a randomized controlled trial of knee arthroscopy for osteoarthritis. *Orthop J Sports Med* 2017;**5**:698439.
24. National Institute for Health and Care Excellence. *Arthroscopic Radiofrequency Chondroplasty for Discrete Chondral Defects of the Knee*. NICE interventional procedures guidance IPG493. London: NICE; 2014.
25. National Institute for Health and Care Excellence. *Arthroscopic Femoro–Acetabular Surgery for Hip Impingement Syndrome*. NICE interventional procedures guidance IPG408. London: NICE; 2011.
26. Palmer AJ, Malak TT, Broomfield J, Holton J, Majkowski L, Thomas GE, *et al.* Past and projected temporal trends in arthroscopic hip surgery in England between 2002 and 2013. *BMJ Open Sport Exerc Med* 2016;**2**:e000082.
27. Pun SY, O'Donnell JM, Kim YJ. Nonarthroplasty hip surgery for early osteoarthritis. *Rheum Dis Clin North Am* 2013; **39**:189–202.
28. Brouwer RW, Raaij van TM, Bierma-Zeinstra SM, Verhagen AP, Jakma TS, Verhaar JA. Osteotomy for treating knee osteoarthritis. *Cochrane Database Syst Rev* 2007:CD004019. <https://doi.org/10.1002/14651858.CD004019>
29. Lerch TD, Steppacher SD, Liechti EF, Tannast M, Siebenrock KA. One-third of hips after periacetabular osteotomy survive 30 years with good clinical results, no progression of arthritis, or conversion to THA. *Clin Orthop Relat Res* 2017;**475**:1154–68.
30. Gomoll AH, Filardo G, de Girolamo L, Espregueira-Mendes J, Marcacci M, Rodkey WG, *et al.* Surgical treatment for early osteoarthritis. Part I: cartilage repair procedures. *Knee Surg Sports Traumatol Arthrosc* 2012;**20**:450–66.
31. Gomoll AH, Filardo G, Almqvist FK, Bugbee WD, Jelic M, Monllau JC, *et al.* Surgical treatment for early osteoarthritis. Part II: allografts and concurrent procedures. *Knee Surg Sports Traumatol Arthrosc* 2012;**20**:468–86.
32. Di Martino A, Kon E, Perdida F, Sessa A, Filardo G, Neri MP, *et al.* Surgical treatment of early knee osteoarthritis with a cell-free osteochondral scaffold: results at 24 months of follow-up. *Injury* 2015;**46**:S33–8.
33. Waugh N, Mistry H, Loveman E, Colquitt J, Royle P, Smith NA, *et al.* *The Use of Allografts in Knee Reconstructions: Systematic Review and Cost-effectiveness Analysis. A Report for the European Society of Sport Traumatology, Knee Surgery and Arthroscopy (ESSKA)*. 2019. URL: https://cdn.ymaws.com/www.esska.org/resource/resmgr/docs/initiatives/allograft/Allografts_report_for_ESSKA_.pdf (accessed 16 December 2019).
34. Palmer JS, Monk AP, Hopewell S, Bayliss LE, Jackson W, Beard DJ, Price AJ. Surgical interventions for early structural knee osteoarthritis. *Cochrane Database Syst Rev* 2016.
35. Luyten FP, Denti M, Filardo G, Kon E, Engebretsen L. Definition and classification of early osteoarthritis of the knee. *Knee Surg Sports Traumatol Arthrosc* 2012;**20**:401–6.
36. Sharma L, Chmiel JS, Almagor O, Felson D, Guermazi A, Roemer F, *et al.* The role of varus and valgus alignment in the initial development of knee cartilage damage by MRI: the MOST study. *Ann Rheum Dis* 2013;**72**:235–40.
37. Felson DT, Niu J, Gross KD, Englund M, Sharma L, Cooke TD, *et al.* Valgus malalignment is a risk factor for lateral knee osteoarthritis incidence and progression: findings from the Multicenter Osteoarthritis Study and the Osteoarthritis Initiative. *Arthritis Rheum* 2013;**65**:355–62.

38. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthritis. *Ann Rheum Dis* 1957;**16**:494–502.
39. Lau BC, Conway D, Mulvihill J, Zhang AL, Feeley BT. Biomechanical consequences of meniscal tear, partial meniscectomy, and meniscal repair in the knee. *JBJS Rev* 2018;**6**:e3.
40. Minkara AA, Westermann RW, Rosneck J, Lynch TS. Systematic review and meta-analysis of outcomes after hip arthroscopy in femoroacetabular impingement. *Am J Sports Med* 2019;**47**:488–500.
41. Matsuda DK, Carlisle JC, Arthurs SC, Wierks CH, Philippon MJ. Comparative systematic review of the open dislocation, mini-open, and arthroscopic surgeries for femoroacetabular impingement. *Arthroscopy* 2011;**27**:252–69.
42. Gao L, Orth P, Cucchiari M, Madry H. Autologous matrix-induced chondrogenesis: a systematic review of the clinical evidence. *Am J Sports Med* 2019;**47**:222–31.
43. National Institute for Health and Care Excellence. *Autologous Chondrocyte Implantation for Treating Symptomatic Articular Cartilage Defects of the Knee*. NICE clinical guideline CG177. London: NICE; 2017.
44. Mistry H, Connock M, Pink J, Shyangdan D, Clar C, Royle P, *et al.* Autologous chondrocyte implantation in the knee: systematic review and economic evaluation. *Health Technol Assess* 2017;**21**:1–294.
45. National Institute for Health and Care Excellence. *Partial Replacement of the Meniscus of the Knee Using a Biodegradable Scaffold*. NICE interventional procedures guidance IPG430. London: NICE; 2012.
46. Chandrasekaran S, Darwish N, Close MR, Lodhia P, Suarez-Ahedo C, Domb BG. Arthroscopic reconstruction of segmental defects of the hip labrum: results in 22 patients with mean 2-year follow-up. *Arthroscopy* 2017;**33**:1685–93.
47. Familiari F, Cinque ME, Chahla J, Godin JA, Olesen ML, Moatshe G, LaPrade RF. Clinical outcomes and failure rates of osteochondral allograft transplantation in the knee: a systematic review. *Am J Sports Med* 2018;**46**:3541–9.
48. De Bari C, Roelofs AJ. Stem cell-based therapeutic strategies for cartilage defects and osteoarthritis. *Curr Opin Pharmacol* 2018;**40**:74–80.
49. Clifford A, O'Connell M, Gabriel S, Miller LE, Block JE. The KineSpring load absorber implant: rationale, design and biomechanical characterization. *J Med Eng Technol* 2011;**35**:65–71.
50. Clarius M, Becker JF, Schmitt H, Seeger JB. The UniSpacer: correcting varus malalignment in medial gonarthrosis. *Int Orthop* 2010;**34**:1175–9.
51. Hirose S, Otsuka H, Morishima T, Sato K. Long-term outcomes of shelf acetabuloplasty for developmental dysplasia of the hip in adults: a minimum 20-year follow-up study. *J Orthop Sci* 2011;**16**:698–703.
52. Shamseer L, Moher D, Clarke M, Ghera D, Liberati A, Petticrew M, *et al.*; PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ* 2015;**350**:g7647.
53. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epidemiol* 2009;**62**:1006–12.
54. Ryd L, Brittberg M, Eriksson K, Jurvelin JS, Lindahl A, Marlovits S, *et al.* Pre-osteoarthritis: definition and diagnosis of an elusive clinical entity. *Cartilage* 2015;**6**:156–65.
55. Siebelt M, Agricola R, Weinans H, Kim YJ. The role of imaging in early hip OA. *Osteoarthritis Cartilage* 2014;**22**:1470–80.
56. Felson DT, Niu J, Guermazi A, Sack B, Aliabadi P. Defining radiographic incidence and progression of knee osteoarthritis: suggested modifications of the Kellgren and Lawrence scale. *Ann Rheum Dis* 2011;**70**:1884–6.
57. Link TM, Steinbach LS, Ghosh S, Ries M, Lu Y, Lane N, Majumdar S. Osteoarthritis: MR imaging findings in different stages of disease and correlation with clinical findings. *Radiology* 2003;**226**:373–81.

58. Williams A, Sharma L, McKenzie CA, Prasad PV, Burstein D. Delayed gadolinium-enhanced magnetic resonance imaging of cartilage in knee osteoarthritis: findings at different radiographic stages of disease and relationship to malalignment. *Arthritis Rheum* 2005;**52**:3528–35.
59. Szebenyi B, Hollander AP, Dieppe P, Quilty B, Duddy J, Clarke S, Kirwan JR. Associations between pain, function, and radiographic features in osteoarthritis of the knee. *Arthritis Rheum* 2006;**54**:230–5.
60. Sheehy L, Culham E, McLean L, Niu J, Lynch J, Segal NA, *et al.* Validity and sensitivity to change of three scales for the radiographic assessment of knee osteoarthritis using images from the Multicenter Osteoarthritis Study (MOST). *Osteoarthritis Cartilage* 2015;**23**:1491–8.
61. Valera M, Ibanez N, Sancho R, Tey M. Reliability of Tonnis classification in early hip arthritis: a useless reference for hip-preserving surgery. *Arch Orthop Trauma Surg* 2016;**136**:27–33.
62. Griffin DR, Dickenson EJ, O'Donnell J, Agricola R, Awan T, Beck M, *et al.* The Warwick Agreement on femoroacetabular impingement syndrome (FAI syndrome): an international consensus statement. *Br J Sports Med* 2016;**50**:1169–76.
63. Hunter DJ, Lo GH, Gale D, Grainger AJ, Guermazi A, Conaghan PG. The reliability of a new scoring system for knee osteoarthritis MRI and the validity of bone marrow lesion assessment: BLOKS (Boston Leeds Osteoarthritis Knee Score). *Ann Rheum Dis* 2008;**67**:206–11.
64. Peterfy CG, Guermazi A, Zaim S, Tirman PF, Miaux Y, White D, *et al.* Whole-organ magnetic resonance imaging score (WORMS) of the knee in osteoarthritis. *Osteoarthritis Cartilage* 2004;**12**:177–90.
65. Hunter DJ, Zhang W, Conaghan PG, Hirko K, Menashe L, Reichmann WM, Losina E. Responsiveness and reliability of MRI in knee osteoarthritis: a meta-analysis of published evidence. *Osteoarthritis Cartilage* 2011;**19**:589–605.
66. Mainil-Varlet P, Aigner T, Brittberg M, Bullough P, Hollander A, Hunziker E, *et al.* Histological assessment of cartilage repair: a report by the Histology Endpoint Committee of the International Cartilage Repair Society (ICRS). *J Bone Joint Surg Am* 2003;**85**:45–57.
67. Outerbridge RE, Dunlop JA. The problem of chondromalacia patellae. *Clin Orthop Relat Res* 1975:177–96.
<https://doi.org/10.1097/00003086-197507000-00024>
68. Bellamy N, Kirwan J, Boers M, Brooks P, Strand V, Tugwell P, *et al.* Recommendations for a core set of outcome measures for future phase III clinical trials in knee, hip, and hand osteoarthritis. Consensus development at OMERACT III. *J Rheumatol* 1997;**24**:799–802.
69. McAlindon TE, Driban JB, Henrotin Y, Hunter DJ, Jiang GL, Skou ST, *et al.* OARSI clinical trials recommendations: design, conduct, and reporting of clinical trials for knee osteoarthritis. *Osteoarthritis Cartilage* 2015;**23**:747–60.
70. Pham T, van der Heijde D, Altman RD, Anderson JJ, Bellamy N, Hochberg M, *et al.* OMERACT-OARSI initiative: osteoarthritis research society international set of responder criteria for osteoarthritis clinical trials revisited. *Osteoarthritis Cartilage* 2004;**12**:389–99.
71. Singh JA, Dohm M, Choong PF. Consensus on draft OMERACT core domains for clinical trials of Total Joint Replacement outcome by orthopaedic surgeons: a report from the International consensus on outcome measures in TJR trials (I-COMiTT) group. *BMC Musculoskelet Disord* 2017;**18**:45.
72. Juni P, Reichenbach S, Dieppe P. Osteoarthritis: rational approach to treating the individual. *Best Pract Res Clin Rheumatol* 2006;**20**:721–40.
73. Juni P, Hari R, Rutjes AW, Fischer R, Silletta MG, Reichenbach S, da Costa BR. Intra-articular corticosteroid for knee osteoarthritis. *Cochrane Database Syst Rev* 2015;**2015**:CD005328.
74. Palmer AJ, Brown CP, McNally EG, Price AJ, Tracey I, Jezzard P, *et al.* Non-invasive imaging of cartilage in early osteoarthritis. *Bone Joint J* 2013;**95-B**:738–46.

75. Guermazi A, Roemer FW, Haugen IK, Crema MD, Hayashi D. MRI-based semiquantitative scoring of joint pathology in osteoarthritis. *Nat Rev Rheumatol* 2013;**9**:236–51.
76. Sterne JAC, Savovic J, Page MJ, Elbers RG, Blencowe NS, Boutron I, *et al.* RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;**366**:l4898.
77. Schunemann H, Hill S, Guyatt G, Akl EA, Ahmed F. The GRADE approach and Bradford Hill's criteria for causation. *J Epidemiol Community Health* 2011;**65**:392–5.
78. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, *et al.*; Cochrane Bias Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;**343**:d5928.
79. Sterne JA, Hernan MA, Reeves BC, Savovic J, Berkman ND, Viswanathan M, *et al.* ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 2016;**355**:i4919.
80. Bisson LJ, Kluczynski MA, Wind WM, Fineberg MS, Bernas GA, Rauh MA, *et al.* Patient outcomes after observation versus debridement of unstable chondral lesions during partial meniscectomy: the chondral lesions and meniscus procedures (ChAMP) randomized controlled trial. *J Bone Joint Surg Am* 2017;**99**:1078–85.
81. Chang RW, Falconer J, Stulberg SD, Arnold WJ, Manheim LM, Dyer AR. A randomized, controlled trial of arthroscopic surgery versus closed-needle joint lavage for patients with osteoarthritis of the knee. *Arthritis Rheum* 1993;**36**:289–96.
82. Forster MC, Straw R. A prospective randomised trial comparing intra-articular Hyalgan injection and arthroscopic washout for knee osteoarthritis. *Knee* 2003;**10**:291–3.
83. Gauffin H, Tagesson S, Meunier A, Magnusson H, Kvist J. Knee arthroscopic surgery is beneficial to middle-aged patients with meniscal symptoms: a prospective, randomised, single-blinded study. *Osteoarthritis Cartilage* 2014;**22**:1808–16.
84. Herrlin S, Hallander M, Wange P, Weidenhielm L, Werner S. Arthroscopic or conservative treatment of degenerative medial meniscal tears: a prospective randomised trial. *Knee Surg Sports Traumatol Arthrosc* 2007;**15**:393–401.
85. Hubbard MJ. Articular debridement versus washout for degeneration of the medial femoral condyle. A five-year study. *J Bone Joint Surg Br* 1996;**78**:217–9.
86. Kalunian KC, Moreland LW, Klashman DJ, Brion PH, Concoff AL, Myers S, *et al.* Visually-guided irrigation in patients with early knee osteoarthritis: a multicenter randomized, controlled trial. *Osteoarthritis Cartilage* 2000;**8**:412–8.
87. Katz JN, Brophy RH, Chaisson CE, de Chaves L, Cole BJ, Dahm DL, *et al.* Surgery versus physical therapy for a meniscal tear and osteoarthritis. *N Engl J Med* 2013;**368**:1675–84.
88. Kirkley A, Birmingham TB, Litchfield RB, Giffin JR, Willits KR, Wong CJ, *et al.* A randomized trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med* 2008;**359**:1097–107.
89. Kise NJ, Risberg MA, Stensrud S, Ranstam J, Engebretsen L, Roos EM. Exercise therapy versus arthroscopic partial meniscectomy for degenerative meniscal tear in middle aged patients: randomised controlled trial with two year follow-up. *Br J Sports Med* 2016;**50**:1473–80.
90. Merchan EC, Galindo E. Arthroscope-guided surgery versus nonoperative treatment for limited degenerative osteoarthritis of the femorotibial joint in patients over 50 years of age: a prospective comparative study. *Arthroscopy* 1993;**9**:663–7.
91. Roos EM, Hare KB, Nielsen SM, Christensen R, Lohmander LS. Better outcome from arthroscopic partial meniscectomy than skin incisions only? A sham-controlled randomised trial in patients aged 35–55 years with knee pain and an MRI-verified meniscal tear. *BMJ Open* 2018;**8**:e019461.
92. Saeed K, Khan SA, Ahmed I. Efficacy of intra articular hyaluronic acid versus arthroscopic debridement in terms of improvement in pain score in Kellgran–Lawrence grading II & III osteoarthritis of knee joint. *Pakistan J Med Health Sci* 2015;**9**:1011–5.

93. Spahn G, Kahl E, Muckley T, Hofmann GO, Klinger HM. Arthroscopic knee chondroplasty using a bipolar radiofrequency-based device compared to mechanical shaver: results of a prospective, randomized, controlled study. *Knee Surg Sports Traumatol Arthrosc* 2008;**16**:565–73.
94. Su X, Li C, Liao W, Liu J, Zhang H, Li J, Li Z. Comparison of arthroscopic and conservative treatments for knee osteoarthritis: a 5-year retrospective comparative study. *Arthroscopy* 2018;**34**:652–9.
95. Van De Graaf VA, Noorduyt JC, Willigenburg NW, Butter IK, De Gast A, Mol BW, *et al.* Effect of early surgery vs physical therapy on knee function among patients with nonobstructive meniscal tears: the ESCAPE randomized clinical trial. *JAMA* 2018;**320**:1328–37.
96. Vermesan D, Prejbeanu R, Laitin S, Damian G, Deleanu B, Abbinante A, *et al.* Arthroscopic debridement compared to intra-articular steroids in treating degenerative medial meniscal tears. *Eur Rev Med Pharmacol Sci* 2013;**17**:3192–6.
97. Yim JH, Seon JK, Song EK, Choi JI, Kim MC, Lee KB, Seo HY. A comparative study of meniscectomy and nonoperative treatment for degenerative horizontal tears of the medial meniscus. *Am J Sports Med* 2013;**41**:1565–70.
98. National Institute for Health and Care Excellence. *Arthroscopic Knee Washout, With or Without Debridement, for the Treatment of Osteoarthritis*. London: NICE; 2007.
99. Rongen JJ, Rovers MM, van Tienen TG, Buma P, Hannink G. Increased risk for knee replacement surgery after arthroscopic surgery for degenerative meniscal tears: a multi-center longitudinal observational study using data from the osteoarthritis initiative. *Osteoarthritis Cartilage* 2017;**25**:23–9.
100. Zikria B, Hafezi-Nejad N, Roemer FW, Guermazi A, Demehri S. Meniscal surgery: risk of radiographic joint space narrowing progression and subsequent knee replacement-data from the osteoarthritis initiative. *Radiology* 2017;**282**:807–16.
101. Zikria B, Hafezi-Nejad N, Wilckens J, Ficke JR, Demehri S. Determinants of knee replacement in subjects with a history of arthroscopy: data from the osteoarthritis initiative. *Eur J Orthop Surg Traumatol* 2016;**26**:665–70.
102. Krych AJ, Thompson M, Knutson Z, Scoon J, Coleman SH. Arthroscopic labral repair versus selective labral debridement in female patients with femoroacetabular impingement: a prospective randomized study. *Arthroscopy* 2013;**29**:46–53.
103. Griffin DR, Dickenson EJ, Wall PDH, Achana F, Donovan JL, Griffin J, *et al.*; FASHIoN Study Group. Hip arthroscopy versus best conservative care for the treatment of femoroacetabular impingement syndrome (UK FASHIoN): a multicentre randomised controlled trial. *Lancet* 2018;**391**:2225–35.
104. Mansell NS, Rhon DI, Meyer J, Slevin JM, Marchant BG. Arthroscopic surgery or physical therapy for patients with femoroacetabular impingement syndrome: a randomized controlled trial with 2-year follow-up. *Am J Sports Med* 2018;**46**:1306–14.
105. Palmer AJR, Ayyar Gupta V, Fernquest S, Rombach I, Dutton SJ, Mansour R, *et al.*; FAIT Study Group. Arthroscopic hip surgery compared with physiotherapy and activity modification for the treatment of symptomatic femoroacetabular impingement: multicentre randomised controlled trial. *BMJ* 2019;**364**:l185.
106. Ayeni OR. Femoroacetabular impingement randomised controlled trial (first): a multicentre randomized controlled trial comparing arthroscopic lavage and arthroscopic osteochondroplasty on patient important outcomes in the treatment of young adult femoroacetabular impingement. *Orthop Proc* 2020;**102-B**:15.
107. Scott BL, Lee CS, Shi LL, Lee MJ, Athiviraham A. Nonoperative management of hip labral tears yields similar total hip arthroplasty conversion rate to arthroscopic treatment. *J Arthroplasty* 2020;**35**:23–7.e1.
108. Heerey J, Risberg MA, Magnus J, Moksnes H, Odegaard T, Crossley K, Kemp JL. Impairment-based rehabilitation following hip arthroscopy: postoperative protocol for the HIP ARThroscopy international randomized controlled trial. *J Orthop Sports Phys Ther* 2018;**48**:336–42.

109. Martin SD, Abraham PF, Varady NH, Nazal MR, Conaway W, Quinlan NJ, Alpaugh K. Hip arthroscopy versus physical therapy for the treatment of symptomatic acetabular labral tears in patients older than 40 years: a randomized controlled trial. *Am J Sports Med* 2021;**49**:1199–208.
110. Glyn-Jones S. *The Arthroscopic Labral Excision or Repair Trial (ALERT)*. 2017. URL: <https://clinicaltrials.gov/study/NCT03257709> (accessed 17 September 2025).
111. Brouwer RW, Bierma-Zeinstra SM, van Raaij TM, Verhaar JA. Osteotomy for medial compartment arthritis of the knee using a closing wedge or an opening wedge controlled by a Puddu plate. A one-year randomised, controlled study. *J Bone Joint Surg Br* 2006;**88**:1454–9.
112. Shaofei Z, Junliang Z, Zhongxin X, Peng H, Xing Z. Comparative analysis of the medium-term effect of the medial unicompartmental arthroplasty through medial approach next to patellar and high tibial osteotomy on medial compartment osteoarthritis of the knee. *Biomed Res* 2017;**28**:3276–80.
113. van der Woude JAD, Wiegant K, van Heerwaarden RJ, Spruijt S, van Roermund PM, Custers RJH, *et al*. Knee joint distraction compared with high tibial osteotomy: a randomized controlled trial. *Knee Surg Sports Traumatol Arthrosc* 2017;**25**:876–86.
114. Van Outeren MV, Waarsing JH, Brouwer RW, Reijman M, Verhaar JAN, Bierma-Zeinstra SMA, *et al*. Is a high tibial osteotomy (HTO) superior to non-surgical treatment or valgus unloader bracing in patients with varus malaligned medial knee osteoarthritis (OA)? A propensity matched study using 2 randomized controlled trial RCT datasets. *Osteoarthritis Cartilage* 2017;**25**:1988–93. <https://doi.org/10.1016/j.joca.2017.09.003>
115. Amin A, Metcalfe A, Scott C, Clement N, Blyth M, Hamilton D, *et al*. Study protocol (NIHR-HTA REF 129820): what is the clinical effectiveness and cost-effectiveness of surgery with medial opening wedge high tibial osteotomy (HTO) compared with non-surgical treatment in the management of osteoarthritis (oa) of the knee in patients younger than 60 years (motion trial)? *Orthop Proc* 2022;**104-B**:12.
116. Brouwer RW, van Raaij TM, Verhaar JA, Coene LN, Bierma-Zeinstra SM. Brace treatment for osteoarthritis of the knee: a prospective randomized multi-centre trial. *Osteoarthritis Cartilage* 2006;**14**:777–83.
117. Kenney R, Domb BG, Lareau J, Cook PC, Giordano BD. Multi-centered comparison of outcomes after hip arthroscopy versus combined hip arthroscopy alone and peri-acetabular osteotomy in patients with acetabular dysplasia. *J Arthroscopic Related Surg* 2018;**34**:e14–e15.
118. Gray BL, Stambough JB, Baca GR, Schoenecker PL, Clohisey JC. Comparison of contemporary periacetabular osteotomy for hip dysplasia with total hip arthroplasty for hip osteoarthritis. *Bone Joint J* 2015;**97-B**:1322–7.
119. Amanatullah DF, Stryker L, Schoenecker P, Taunton MJ, Clohisey JC, Trousdale RT, Sierra RJ. Similar clinical outcomes for THAs with and without prior periacetabular osteotomy. *Clin Orthop Relat Res* 2015;**473**:685–91.
120. Fukui K, Kaneuji A, Sugimori T, Ichiseki T, Matsumoto T. Does rotational acetabular osteotomy affect subsequent total hip arthroplasty? *Arch Orthop Trauma Surg* 2015;**135**:407–15.
121. Hashemi-Nejad A, Haddad FS, Tong KM, Muirhead-Allwood SK, Catterall A. Does Chiari osteotomy compromise subsequent total hip arthroplasty? *J Arthroplasty* 2002;**17**:731–9.
122. Ito H, Takatori Y, Moro T, Oshima H, Oka H, Tanaka S. Total hip arthroplasty after rotational acetabular osteotomy. *J Arthroplasty* 2015;**30**:403–6.
123. Migaud H, Putman S, Berton C, Lefevre C, Hutten D, Argenson JN, Gaucher F. Does prior conservative surgery affect survivorship and functional outcome in total hip arthroplasty for congenital dislocation of the hip? A case-control study in 159 hips. *Orthop Traumatol Surg Res* 2014;**100**:733–7.
124. Osawa Y, Hasegawa Y, Seki T, Takegami Y, Amano T, Ishiguro N. Patient-reported outcomes in patients who undergo total hip arthroplasty after periacetabular osteotomy. *J Orthop Sci* 2018;**23**:346–9.
125. Stambough JB, Clohisey JC, Baca GR, Zaltz I, Trousdale R, Millis M, *et al*. Does previous pelvic osteotomy compromise the results of periacetabular osteotomy surgery? *Clin Orthop Relat Res* 2015;**473**:1417–24.

126. Tokunaga K, Aslam N, Zdero R, Schemitsch EH, Waddell JP. Effect of prior Salter or Chiari osteotomy on THA with developmental hip dysplasia. *Clin Orthop Relat Res* 2011;**469**:237–43.
127. Yacovelli S, Abdelaal M, Fillingham Y, Sutton R, Madding R, Parvizi J. Prior pelvic osteotomy affects the outcome of subsequent total hip arthroplasty. *J Arthroplasty* 2021;**36**:600–4.
128. Yuasa T, Maezawa K, Nozawa M, Kaneko K. Total hip arthroplasty after previous rotational acetabular osteotomy. *Eur J Orthop Surg Traumatol* 2015;**25**:1057–60.
129. Beck EC, Gowd AK, Paul K, Chahla J, Marquez-Lara AJ, Rasio J, *et al.* Pelvic osteotomies for acetabular dysplasia: are there outcomes, survivorship and complication differences between different osteotomy techniques? *J Hip Preserv Surg* 2021;**7**:764–76.
130. Sohatee MA, Ali M, Khanduja V, Malviya A. Does hip preservation surgery prevent arthroplasty? Quantifying the rate of conversion to arthroplasty following hip preservation surgery. *J Hip Preserv Surg* 2020;**7**:168–82.
131. Reimer LCU, Jakobsen SS, Mortensen L, Dalgas U, Jacobsen JS, Soballe K, *et al.* Efficacy of periacetabular osteotomy followed by progressive resistance training compared to progressive resistance training as non-surgical treatment in patients with hip dysplasia (PreserveHip) – a protocol for a randomised controlled trial. *BMJ Open* 2019;**9**:e032782.
132. Kim MS, Chun CH, Wang JH, Kim JG, Kang SB, Yoo JD, *et al.* Microfractures versus a porcine-derived collagen-augmented chondrogenesis technique for treating knee cartilage defects: a multicenter randomized controlled trial. *Arthroscopy* 2020;**36**:1612–24.
133. Gudas R, Gudaite A, Pocius A, Gudiene A, Cekanauskas E, Monastyreckiene E, Basevicius A. Ten-year follow-up of a prospective, randomized clinical study of mosaic osteochondral autologous transplantation versus microfracture for the treatment of osteochondral defects in the knee joint of athletes. *Am J Sports Med* 2012;**40**:2499–508.
134. Gudas R, Stankevicius E, Monastyreckiene E, Pranys D, Kalesinskas RJ. Osteochondral autologous transplantation versus microfracture for the treatment of articular cartilage defects in the knee joint in athletes. *Knee Surg Sports Traumatol Arthrosc* 2006;**14**:834–42.
135. Ulstein S, Aroen A, Rotterud JH, Loken S, Engebretsen L, Heir S. Microfracture technique versus osteochondral autologous transplantation mosaicplasty in patients with articular chondral lesions of the knee: a prospective randomized trial with long-term follow-up. *Knee Surg Sports Traumatol Arthrosc* 2014;**22**:1207–15.
136. Lim HC, Bae JH, Song SH, Park YE, Kim SJ. Current treatments of isolated articular cartilage lesions of the knee achieve similar outcomes. *Clin Orthop Relat Res* 2012;**470**:2261–7.
137. Solheim E, Hegna J, Strand T, Harlem T, Inderhaug E. Randomized study of long-term (15–17 years) outcome after microfracture versus mosaicplasty in knee articular cartilage defects. *Am J Sports Med* 2018;**46**:826–31.
138. Kon E, Filardo G, Brittberg M, Busacca M, Condello V, Engebretsen L, *et al.* A multilayer biomaterial for osteochondral regeneration shows superiority vs microfractures for the treatment of osteochondral lesions in a multicentre randomized trial at 2 years. *Knee Surg Sports Traumatol Arthrosc* 2018;**26**:2704–15.
139. Anders S, Volz M, Frick H, Gellissen J. A randomized, controlled trial comparing autologous matrix-induced chondrogenesis [AMIC(R)] to microfracture: analysis of 1- and 2-year follow-up data of 2 centers. *Open Orthop J* 2013;**7**:133–43.
140. Volz M, Schaumburger J, Frick H, Grifka J, Anders S. A randomized controlled trial demonstrating sustained benefit of Autologous Matrix-Induced Chondrogenesis over microfracture at five years. *Int Orthop* 2017;**41**:797–804.
141. Shive MS, Stanish WD, McCormack R, Forriol F, Mohtadi N, Pelet S, *et al.* BST-CarGel(R) treatment maintains cartilage repair superiority over microfracture at 5 years in a multicenter randomized controlled trial. *Cartilage* 2015;**6**:62–72.

142. Fossum V, Hansen AK, Wilsgaard T, Knutsen G. Collagen-covered autologous chondrocyte implantation versus autologous matrix-induced chondrogenesis: a randomized trial comparing 2 methods for repair of cartilage defects of the knee. *Orthop J Sports Med* 2019;**7**:2325967119868212.
143. Knutsen G, Drogset JO, Engebretsen L, Grontvedt T, Isaksen V, Ludvigsen TC, *et al.* A randomized trial comparing autologous chondrocyte implantation with microfracture. Findings at five years. *J Bone Joint Surg Am* 2007;**89**:2105–12.
144. Knutsen G, Drogset JO, Engebretsen L, Grontvedt T, Ludvigsen TC, Loken S, *et al.* A randomized multicenter trial comparing autologous chondrocyte implantation with microfracture: long-term follow-up at 14 to 15 years. *J Bone Joint Surg Am* 2016;**98**:1332–9.
145. Knutsen G, Engebretsen L, Ludvigsen TC, Drogset JO, Grontvedt T, Solheim E, *et al.* Autologous chondrocyte implantation compared with microfracture in the knee. A randomized trial. *J Bone Joint Surg Am* 2004;**86**:455–64.
146. Saris DB, Vanlauwe J, Victor J, Haspl M, Bohnsack M, Fortems Y, *et al.* Characterized chondrocyte implantation results in better structural repair when treating symptomatic cartilage defects of the knee in a randomized controlled trial versus microfracture. *Am J Sports Med* 2008;**36**:235–46.
147. Van Assche D, Staes F, Van Caspel D, Vanlauwe J, Bellemans J, Saris DB, Luyten FP. Autologous chondrocyte implantation versus microfracture for knee cartilage injury: a prospective randomized trial, with 2-year follow-up. *Knee Surg Sports Traumatol Arthrosc* 2010;**18**:486–95.
148. Vanlauwe J, Saris DB, Victor J, Almqvist KF, Bellemans J, Luyten FP; TIG/ACT/01/2000&EXT Study Group. Five-year outcome of characterized chondrocyte implantation versus microfracture for symptomatic cartilage defects of the knee: early treatment matters. *Am J Sports Med* 2011;**39**:2566–74.
149. Saris D, Price A, Widuchowski W, Bertrand-Marchand M, Caron J, Drogset JO, *et al.*; SUMMIT study group. Matrix-applied characterized autologous cultured chondrocytes versus microfracture: two-year follow-up of a prospective randomized trial. *Am J Sports Med* 2014;**42**:1384–94.
150. Brittberg M, Recker D, Ilgenfritz J, Saris DBF, Group SES. Matrix-applied characterized autologous cultured chondrocytes versus microfracture: five-year follow-up of a prospective randomized trial. *Am J Sports Med* 2018;**46**:1343–51.
151. Crawford DC, DeBerardino TM, Williams RJ 3rd. NeoCart, an autologous cartilage tissue implant, compared with microfracture for treatment of distal femoral cartilage lesions: an FDA phase-II prospective, randomized clinical trial after two years. *J Bone Joint Surg Am* 2012;**94**:979–89.
152. Basad E, Ishaque B, Bachmann G, Sturz H, Steinmeyer J. Matrix-induced autologous chondrocyte implantation versus microfracture in the treatment of cartilage defects of the knee: a 2-year randomised study. *Knee Surg Sports Traumatol Arthrosc* 2010;**18**:519–27.
153. Clave A, Potel JF, Servien E, Neyret P, Dubrana F, Stindel E. Third-generation autologous chondrocyte implantation versus mosaicplasty for knee cartilage injury: 2-year randomized trial. *J Orthop Res* 2016;**34**:658–65.
154. Niemeyer P, Laute V, Zinser W, Becher C, Kolombe T, Fay J, *et al.* A prospective, randomized, open-label, multicenter, phase III noninferiority trial to compare the clinical efficacy of matrix-associated autologous chondrocyte implantation with spheroid technology versus arthroscopic microfracture for cartilage defects of the knee. *Orthop J Sports Med* 2019;**7**:2325967119854442.
155. Whitehouse M, Marques E, Culliford L, Howells N, Burston A, Judge A, *et al.* A Randomised Controlled Trial of Scaffold InSertion and Microfracture Compared to Microfracture Alone for the Treatment of Chondral or Osteochondral Defects of the Knee: The SSMIC Study. NIHR; 2019. URL: <https://fundingawards.nihr.ac.uk/award/NIHR127849> (accessed 17 September 2025).
156. National Institute for Health and Care Excellence. *Autologous Chondrocyte Implantation for Treating Symptomatic Articular Cartilage Defects of the Knee*. 2017. URL: www.nice.org.uk/guidance/ta477 (accessed 17 September 2025).

157. Chaharbakhshi EO, Hartigan DE, Spencer JD, Perets I, Lall AC, Domb BG. Do larger acetabular chondral defects portend inferior outcomes in patients undergoing arthroscopic acetabular microfracture? A matched-controlled study. *Arthroscopy* 2019;**35**:2037–47.
158. de Girolamo L, Jannelli E, Fioruzzi A, Fontana A. Acetabular chondral lesions associated with femoroacetabular impingement treated by autologous matrix-induced chondrogenesis or microfracture: a comparative study at 8-year follow-up. *Arthroscopy* 2018;**34**:3012–23.
159. Domb BG, Gupta A, Dunne KF, Gui C, Chandrasekaran S, Lodhia P. Microfracture in the hip: results of a matched-cohort controlled study with 2-year follow-up. *Am J Sports Med* 2015;**43**:1865–74.
160. Fontana A, de Girolamo L. Sustained five-year benefit of autologous matrix-induced chondrogenesis for femoral acetabular impingement-induced chondral lesions compared with microfracture treatment. *Bone Joint J* 2015;**97-B**:628–35.
161. Hartigan DE, Perets I, Chaharbakhshi EO, Walsh JP, Yuen LC, Domb BG. Outcomes of femoral head marrow stimulation techniques at minimum 2-year follow-up. *Orthopedics* 2018;**41**:e70–6.
162. Hevesi M, Bernard C, Hartigan DE, Levy BA, Domb BG, Krych AJ. Is microfracture necessary? Acetabular chondrolabral debridement/abrasion demonstrates similar outcomes and survival to microfracture in hip arthroscopy: a multicenter analysis. *Am J Sports Med* 2019;**47**:1670–8.
163. John R, Ma J, Wong I. Better clinicoradiological results of BST-CarGel treatment in cartilage repair compared with microfracture in acetabular chondral defects at 2 years. *Am J Sports Med* 2020;**48**:1961–6.
164. Kester BS, Begly JP, Capogna B, Chenard K, Youm T. Four-year outcomes following arthroscopic microfracture of the hip in patients with advanced chondral lesions. *Bull Hosp Jt Dis (2013)* 2020;**78**:123–30.
165. Mancini D, Fontana A. Five-year results of arthroscopic techniques for the treatment of acetabular chondral lesions in femoroacetabular impingement. *Int Orthop* 2014;**38**:2057–64.
166. Rivera E, Seijas R, Rubio M, Garcia-Balletbo M, Vilar JM, Boada PL, Cugat R. Outcomes at 2-years follow-up after hip arthroscopy combining bone marrow concentrate. *J Invest Surg* 2020;**33**:655–63.
167. Skou ST, Roos EM, Laursen MB, Rathleff MS, Arendt-Nielsen L, Simonsen O, Rasmussen S. A randomized, controlled trial of total knee replacement. *N Engl J Med* 2015;**373**:1597–606.
168. Cushnaghan J, Bennett J, Reading I, Croft P, Byng P, Cox K, *et al.* Long-term outcome following total knee arthroplasty: a controlled longitudinal study. *Ann Rheum Dis* 2009;**68**:642–7.
169. Pandit H, Gulati A, Jenkins C, Barker K, Price AJ, Dodd CA, Murray DW. Unicompartmental knee replacement for patients with partial thickness cartilage loss in the affected compartment. *Knee* 2011;**18**:168–71.
170. Keurentjes JC, Fiocco M, So-Osman C, Onstenk R, Koopman-Van Gemert AW, Poll RG, *et al.* Patients with severe radiographic osteoarthritis have a better prognosis in physical functioning after hip and knee replacement: a cohort-study. *PLOS ONE* 2013;**8**:e59500.
171. Riis A, Rathleff MS, Jensen MB, Simonsen O. Low grading of the severity of knee osteoarthritis pre-operatively is associated with a lower functional level after total knee replacement: a prospective cohort study with 12 months' follow-up. *Bone Joint J* 2014;**96-B**:1498–502.
172. Maier MW, Kuhs F, Streit MR, Schuhmacher P, Walker T, Ewerbeck V, Gotterbarm T. Unicompartmental knee arthroplasty in patients with full versus partial thickness cartilage loss (PTCL): equal in clinical outcome but with higher reoperation rate for patients with PTCL. *Arch Orthop Trauma Surg* 2015;**135**:1169–75.
173. Tilbury C, Holtslag MJ, Tordoir RL, Leichtenberg CS, Verdegaal SH, Kroon HM, *et al.* Outcome of total hip arthroplasty, but not of total knee arthroplasty, is related to the preoperative radiographic severity of osteoarthritis. A prospective cohort study of 573 patients. *Acta Orthop* 2016;**87**:67–71.
174. Hamilton TW, Pandit HG, Inabathula A, Ostlere SJ, Jenkins C, Mellon SJ, *et al.* Unsatisfactory outcomes following unicompartmental knee arthroplasty in patients with partial thickness cartilage loss: a medium-term follow-up. *Bone Joint J* 2017;**99-B**:475–82.

175. Hoorntje A, Witjes S, Koenraadt KLM, Aarts R, Weert T, van Geenen RCI. More severe preoperative Kellgren-Lawrence grades of knee osteoarthritis were partially associated with better postoperative patient-reported outcomes in TKA patients. *J Knee Surg* 2019;**32**:211–7.
176. van de Water RB, Leichtenberg CS, Nelissen R, Kroon HM, Kaptijn HH, Onstenk R, *et al*. Preoperative radiographic osteoarthritis severity modifies the effect of preoperative pain on pain/function after total knee arthroplasty: results at 1 and 2 years postoperatively. *J Bone Joint Surg Am* 2019;**101**:879–87.
177. Lange JK, Yang HY, Collins JE, Losina E, Katz JN. Association between preoperative radiographic severity of osteoarthritis and patient-reported outcomes of total knee replacement. *JB JS Open Access* 2020;**5**:e19.00073.
178. Rehman Y, Lindberg MF, Arnljot K, Gay CL, Lerdal A, Aamodt A. More severe radiographic osteoarthritis is associated with increased improvement in patients' health state following a total knee arthroplasty. *J Arthroplasty* 2020;**35**:3131–7.
179. Kim CH, Kim JW. Periacetabular osteotomy vs. total hip arthroplasty in young active patients with dysplastic hip: systematic review and meta-analysis. *Orthop Traumatol Surg Res* 2020;**106**:1545–51.
180. Al-Amiry BS, Gaber JF, Kadum BK, Brismar TB, Sayed-Noor AS. The influence of radiological severity and symptom duration of osteoarthritis on postoperative outcome after total hip arthroplasty: a prospective cohort study. *J Arthroplasty* 2018;**33**:436–40.
181. Judge A, Cooper C, Arden NK, Williams S, Hobbs N, Dixon D, *et al*. Pre-operative expectation predicts 12-month post-operative outcome among patients undergoing primary total hip replacement in European orthopaedic centres. *Osteoarthritis Cartilage* 2011;**19**:659–67.
182. Judge A, Cooper C, Williams S, Dreinhoefer K, Dieppe P. Patient-reported outcomes one year after primary hip replacement in a European Collaborative Cohort. *Arthritis Care Res (Hoboken)* 2010;**62**:480–8.
183. Cushnaghan J, Coggon D, Reading I, Croft P, Byng P, Cox K, *et al*. Long-term outcome following total hip arthroplasty: a controlled longitudinal study. *Arthritis Care Res* 2007;**57**:1375–80.
184. Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. *Med Decis Making* 1993;**13**:322–38.
185. UK NCGC. *Osteoarthritis: Care and Management in Adults*. London: NICE; 2014.
186. Wai EK, Kreder HJ, Williams JI. Arthroscopic debridement of the knee for osteoarthritis in patients fifty years of age or older: utilization and outcomes in the Province of Ontario. *J Bone Joint Surg Am* 2002;**84**:17–22.
187. Guyot P, Ades A, Ouwers MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan–Meier survival curves. *BMC Med Res Methodol* 2012;**12**:1–13.
188. Gu A, Malahias MA, Cohen JS, Richardson SS, Stake S, Blevins JL, Sculco PK. Prior knee arthroscopy is associated with increased risk of revision after total knee arthroplasty. *J Arthroplasty* 2020;**35**:100–4.
189. Briggs A, Sculpher M, Claxton K. *Decision Modelling for Health Economic Evaluation*. Oxford: Oxford University Press; 2006.
190. Office for National Statistics. *National Life Tables, United Kingdom, 2018–2020*. 2021 [cited 2025 5 April 2025]. URL: www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables (accessed 17 September 2025).
191. Bolarinwa SA, Aryee JN, Labaran LA, Werner BC, Browne JA. Does arthroscopic repair of femoroacetabular impingement pathology affect clinical outcomes after ipsilateral total hip arthroplasty? *Hip Pelvis* 2020;**32**:35–41.
192. Malahias MA, Alexiades M. The clinical outcome of chondrolabral-preserving arthroscopic acetabuloplasty for pincer-or mixed-type femoroacetabular impingement: a systematic review. *Musculoskelet Surg* 2019;**103**:207–14.

193. Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997;**35**:1095–108.
194. National Institute for Health and Care Excellence. *Guide to the Methods of Technology Appraisal*. London: NICE; 2013.
195. Barton GR, Sach TH, Jenkinson C, Avery AJ, Doherty M, Muir KR. Do estimates of cost-utility based on the EQ-5D differ from those based on the mapping of utility scores? *Health Qual Life Outcomes* 2008;**6**:1–11.
196. Dakin H, Abel L, Burns R, Yang Y. Review and critical appraisal of studies mapping from quality of life or clinical measures to EQ-5D: an online database and application of the MAPS statement. *Health Qual Life Outcomes* 2018;**16**:1–9.
197. Kiran A, Hunter DJ, Judge A, Field RE, Javaid MK, Cooper C, Arden NK; Oxford NDORMS Musculoskeletal Epidemiology Unit Writing Committee. A novel methodological approach for measuring symptomatic change following total joint arthroplasty. *J Arthroplasty* 2014;**29**:2140–5.
198. Curtis LA, Burns A. *Unit Costs of Health and Social Care 2015*. Unit Costs of Health and Social Care. Report no 3. Kent: Personal Social Services Research Unit, University of Kent, UK; 2020.
199. Mistry H, Metcalfe A, Smith N, Loveman E, Colquitt J, Royle P, Waugh N. The cost-effectiveness of osteochondral allograft transplantation in the knee. *Knee Surg Sports Traumatol Arthrosc* 2019;**27**:1739–53.
200. Burn E, Edwards CJ, Murray DW, Silman A, Cooper C, Arden NK, *et al*. Trends and determinants of length of stay and hospital reimbursement following knee and hip replacement: evidence from linked primary care and NHS hospital records from 1997 to 2014. *BMJ Open* 2018;**8**:e019146.
201. Morgan OJ, Hillstrom HJ, Ellis SJ, Golightly YM, Russell R, Hannan MT, *et al*. Osteoarthritis in England: incidence trends from national health service hospital episode statistics. *ACR Open Rheumatology* 2019;**1**:493–8.
202. Drummond M, Drummond MF, McGuire A. *Economic Evaluation in Health Care: Merging Theory with Practice*. Oxford: Oxford University Press; 2001.
203. Fox M, Mealing S, Anderson R, Dean J, Stein K, Price A, Taylor RS. The clinical effectiveness and cost-effectiveness of cardiac resynchronisation (biventricular pacing) for heart failure: systematic review and economic model. *Health Technol Assess* 2007;**11**:iii–v, ix.
204. Kang S. Assessing responsiveness of the EQ-5D-3L, the Oxford Hip Score, and the Oxford Knee Score in the NHS patient-reported outcome measures. *J Orthop Surg Res* 2021;**16**:1–12.
205. National Institute for Health and Care Excellence. *Guide to the Methods of Technology Appraisal 2013*. London: NICE; 2013.

Appendix 1

Score systematic review – Search strategies for MEDLINE (Ovid)

1. Non-regenerative surgical procedures used to treat early osteoarthritis in the knee.

1. *knee/su
2. *Knee Joint/su
3. *knee osteoarthritis/su
4. *Patellofemoral Joint/su
5. *Patella/su
6. (knee or knees or patella* or patellofemoral or condyl* or tibia* or fibula*).tw,kf.
7. 1 or 2 or 3 or 4 or 5 or 6
8. *Arthroscopy/
9. arthroscop*.tw,kf.
10. chondroplast*.tw,kf.
11. (arthroscopic adj2 debridement).tw,kf.
12. (chondral adj2 debridement).tw,kf.
13. (BioPoly or Episealer or Episurf or Hemicap or Unicap or Miaoregen or Agili-C or Trufit).tw,kf.
14. 8 or 9 or 10 or 11 or 12 or 13
15. (meniscal adj2 implant*).tw,kf.
16. (meniscal adj2 scaffold*).tw,kf.
17. 14 or 15 or 16
18. exp clinical study/ or comparative study/ or meta-analysis/ or multicenter study/ or "systematic review"/
19. (randomized controlled trial or controlled clinical trial or clinical trial or systematic review or meta-analysis or comparative study or multicenter study).pt.
20. (compare* or comparison or comparative or versus or matched or group* or divided or cohort).tw,kf.
21. (random* or trial* or blind* or placebo or control* or systematic review or meta-analysis).kf,tw.
22. 18 or 19 or 20 or 21
23. (editorial or comment or letter or case reports).pt.
24. animals/
25. humans/
26. 24 not 25
27. 7 and 17 and 22
28. limit 27 to english language
29. 28 not 23
30. 29 not 26
31. (meniscectomy or osteonecrosis or fracture* or osteochondritis dissecans or meniscal tear* or anterior cruciate ligament or ACL).ti.
32. 30 not 31

2. Non-regenerative surgical procedures used to treat early osteoarthritis in the hip.

- 1 Hip Joint/ or Hip/ or Osteoarthritis, Hip/
- 2 Femoracetabular Impingement/
- 3 Acetabulum/
- 4 Hip Dislocation, Congenital/
- 5 Hip Dislocation/
- 6 (hip or hips or acetabular or acetabulum or femoracetabular or FAI or labrum or labral or cam or pincer).tw,kw.
- 7 1 or 2 or 3 or 4 or 5 or 6
- 8 Acetabuloplasty/

- 9 Acetabuloplast*.tw,kf.
- 10 Arthroscopy/
- 11 Arthroscop*.tw,kf.
- 12 osteochondroplast*.tw,kf.
- 13 Debridement/ or debridement.tw,kf.
- 14 rim trimming.tw,kf.
- 15 (labral adj2 (deconstruction or repair or surgery or graf*)).tw,kw.
- 16 mini-open.tw,kf.
- 17 (open adj2 surgery).tw,kf.
- 18 Ganz trochanteric flip.tw,kf.
- 19 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18
- 20 7 and 19
- 21 limit 20 to english language
- 22 exp clinical study/ or comparative study/ or meta-analysis/ or multicenter study/ or "systematic review"/
- 23 (randomized controlled trial or controlled clinical trial or clinical trial or systematic review or meta-analysis or comparative study or multicenter study).pt.
- 24 cohort studies/ or prospective studies/ or retrospective studies/
- 25 (compare* or comparison or comparative or versus or matched or group* or divided or cohort).tw,kf.
- 26 (prospective or retrospective or case series).tw,kf.
- 27 (random* or trial* or blind* or placebo or control* or systematic review or meta-analysis).kf,tw.
- 28 22 or 23 or 24 or 25 or 26 or 27
- 29 (editorial or comment or letter or case reports).pt.
- 30 animals/
- 31 humans/
- 32 30 not 31
- 33 21 and 28
- 34 33 not 29
- 35 34 not 32

3. Load-modifying procedures used to treat early osteoarthritis in the knee

1. exp *Knee Joint/su [Surgery]
2. *Osteoarthritis, Knee/su [Surgery]
3. exp *Knee/su [Surgery]
4. exp *Tibia/su [Surgery]
5. (knee* or tibia*).tw,kw.
6. 1 or 2 or 3 or 4 or 5
7. Osteotomy/
8. osteotom*.tw,kw.
9. 7 or 8
10. 6 and 9
11. (knee adj5 joint adj5 distraction).tw,kw.
12. Distal femoral osteotom*.tw,kw.
13. high tibia* osteotom*.tw,kw.
14. (Opty-line or kinespring or Moximed or Tomofix or (atlas* adj2 knee)).tw,kw.
15. ((interpositional adj3 (implant* or device*)) and knee).tw,kw.
16. 11 or 12 or 13 or 14 or 15
17. 10 or 16
18. (case reports or editorial or comment or letter).pt.
19. 17 not 18
20. exp clinical study/or comparative study/or meta-analysis/or multicenter study/or 'systematic review'/
21. (randomized controlled trial or controlled clinical trial or clinical trial or systematic review or meta-analysis or comparative study or multicenter study).pt.

22. (compare* or comparison or comparative or versus or matched or group* or divided or cohort).tw,kw.
23. (random* or trial* or blind* or control* or systematic review or meta-analysis).kw,tw.
24. 20 or 21 or 22 or 23
25. 19 and 24
26. Animals/
27. Humans/
28. 26 not 27
29. 25 not 28
30. limit 29 to english language

4. Load-modifying procedures used to treat early osteoarthritis in the hip

1. Osteotomy/
2. *Acetabuloplasty/ or *bone remodeling/
3. (osteotom* or shelf acetabuloplasty or cheilectomy or osteoplasty).tw,kw.
4. 1 or 2 or 3
5. Hip Joint/ or Hip/ or Osteoarthritis, Hip/
6. Femoracetabular Impingement/su [Surgery]
7. (hip or hips or femoracetabular impingement or shelf acetabuloplasty).tw,kw.
8. (pelvic or femoral head).ti.
9. 5 or 6 or 7 or 8
10. 4 and 9
11. (case reports or editorial or comment or letter).pt.
12. 10 not 11
13. exp clinical study/ or comparative study/ or meta-analysis/ or multicenter study/ or "systematic review"/ or Observational Study/
14. (randomized controlled trial or controlled clinical trial or clinical trial or systematic review or meta-analysis or comparative study or multicenter study).pt.
15. (compare* or comparison or comparative or versus or matched or group* or divided or cohort or observational).tw,kw.
16. (random* or trial* or blind* or control* or prospective* or systematic review or meta-analysis).kw,tw.
17. 13 or 14 or 15 or 16
18. 12 and 17
19. Animals/
20. Humans/
21. 19 not 20
22. 18 not 21

5. Biological reconstruction used to treat early osteoarthritis in the knee

1. exp *Chondrocytes/tr [Transplantation]
2. exp *Cartilage, Articular/su, tr [Surgery, Transplantation]
3. exp *Transplantation, Autologous/
4. (MACI or MACT or ACI).tw,kf.
5. (chondrocyte* adj4 (implant* or transplant* or matrix or characteri*)).tw,kf.
6. (autologous adj3 (implant* or transplant* or chondrocyte* or cartilage)).tw,kf.
7. (matrix adj3 autologous adj3 chondrocyte*).tw,kf.
8. (cartilage* adj2 (transplant* or implant*)).tw,kf.
9. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8
10. allogenic chondrocyte*.tw,kf.
11. exp *Allografts/
12. allograft*.tw,kw.
13. mosaicplasty.kf,tw.

14. (microfracture or AMIC or autologous matrix-induced chondrogenesis).tw,kf.
15. *Platelet-Rich Plasma/ or platelet rich plasma.tw,kf.
16. *Stem Cells/ or *stem cell transplantation/ or *mesenchymal stem cell transplantation/
17. (bone marrow or adipose-derived or stem cell* or (mesenchymal adj2 cell*)).ti.
18. (agili-C or cartione or chondroelect or maioregen or novocart or OsCell or syngenit or tigenix or vericel).mp.
19. (osteochondral adj3 (scaffold* or repair* or defect* or regeneration)).tw,kf.
20. 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19
21. Osteoarthritis, Knee/ or Knee/ or Knee Joint/
22. (knee* or patella* or condyle* or patellofemoral).tw,kf.
23. 21 or 22
24. 9 or 20
25. 23 and 24
26. exp clinical study/ or comparative study/ or meta-analysis/ or multicenter study/ or "systematic review"/
27. (randomized controlled trial or controlled clinical trial or clinical trial or systematic review or meta-analysis or comparative study or multicenter study).pt.
28. (compare* or comparison or comparative or versus or matched or group* or divided or cohort).tw,kf.
29. (random* or trial* or blind* or placebo or control* or systematic review or meta-analysis).kf,tw.
30. 26 or 27 or 28 or 29
31. 25 and 30
32. limit 31 to english language
33. limit 32 to yr="1980 -Current"
34. animals/
35. Humans/
36. 34 not 35
37. 33 not 36
38. (editorial or comment or letter or case reports).pt.
39. 37 not 38

6. Biological reconstruction used to treat early osteoarthritis in the hip

1. *chondrocyte/ or *chondrocyte implantation/ or *chondrocyte implant/
2. *articular cartilage/su [Surgery]
3. (MACI or MACT or ACI).tw,kw.
4. allogenic chondrocyte*.tw,kw.
5. (autologous adj2 (implant* or transplant* or chondrocyte* or cartilage or reconstruction)).tw,kw.
6. autologous chondrocyte* implant*.tw,kw.
7. (autologous matrix-induced chondrogenesis or AMIC).tw,kw.
8. (matrix* adj3 chondrocyte* adj3 Implant*).tw,kw.
9. (osteochondral adj3 (repair* or defect* or regenerat* or scaffold or patch or fleece or graft*)).tw,kw.
10. *allograft/
11. (allograft* or mosaicplasty or microfracture*).tw,kw.
12. *thrombocyte rich plasma/
13. platelet rich plasma.tw,kw.
14. exp *mesenchymal stem cell/
15. stem cell/ or adipose-derived stem cell/
16. (bone marrow or adipose-derived or stem cell* or (mesenchymal adj2 cell*)).tw,kw.
17. (agili-c or maioregen).tw,kw.
18. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17
19. acetabulum/su [Surgery]
20. hip/su [Surgery]
21. hip osteoarthritis/
22. femoroacetabular impingement/su [Surgery]
23. femoral head/

24. (hip or hips or acetabular or acetabulum or femoracetabular or FAI or labrum or labral or femoral head).tw,kw.
25. 19 or 20 or 21 or 22 or 23 or 24
26. 18 and 25
27. exp clinical study/ or comparative study/ or meta-analysis/ or multicenter study/ or "systematic review"/
28. cohort analysis/ or prospective study/ or retrospective study/
29. (compare* or comparison or comparative or versus or matched or group* or divided or cohort or case series).tw,kw.
30. (random* or trial* or blind* or placebo or control* or systematic review or meta-analysis).kw,tw.
31. 27 or 28 or 29 or 30
32. 26 and 31
33. (letter or editorial).pt.
34. case report/
35. 33 or 34
36. animal/
37. human/
38. 36 not 37
39. conference abstract.pt.
40. 33 or 34 or 39
41. 32 not 40
42. 41 not 38
43. (osteonecrosis or osteochondritis dissecans).m_titl.
44. 42 not 43
45. limit 44 to (english language and yr="1980 -Current")

9. Knee arthroplasty used to treat early osteoarthritis in the knee

1. Knee Joint/su [Surgery]
2. Knee/su [Surgery]
3. Patellofemoral Joint/su [Surgery]
4. Patella/su [Surgery]
5. Knee Prosthesis/
6. Arthroplasty, Replacement, Knee/
7. 1 or 2 or 3 or 4 or 5 or 6
8. (total knee adj2 (replacement or arthroplast*)).tw,kf.
9. (partial knee adj2 (replacement or arthroplast*)).tw,kf.
10. (unicompartmental knee adj2 (replacement or arthroplast*)).tw,kf.
11. (Patellofemoral adj2 (Replacement or arthroplast*)).tw,kf.
12. (TKR or UKR or UKA or TKA).tw,kf.
13. 8 or 9 or 10 or 11 or 12
14. 7 and 13
15. randomized controlled trial.pt.
16. (random* or blind*).tw,kf.
17. (systematic review or meta-analysis).pt.
18. (systematic review or meta-analysis).tw,kf.
19. 15 or 16 or 17 or 18
20. 14 and 19
21. ((knee or patellofemoral) adj2 (replacement* or arthroplast*)).m_titl.
22. (osteoarthritis and (surgery or surgical)).tw,kf.
23. 19 and 21 and 22
24. 20 or 23
25. (editorial or comment or letter).pt.
26. 24 not 25
27. limit 26 to english language
28. limit 27 to yr="2000 -Current"

10. Hip arthroplasty used to treat early osteoarthritis in the hip

1. exp Arthroplasty, Replacement, Hip/
2. Hip Joint/
3. Joint Prosthesis/
4. hip prosthesis/ or metal-on-metal joint prostheses/
5. (hip adj2 (replacement or arthroplasty)).tw,kf.
6. (THA or THR).tw,kf.
7. metal-on-metal.tw,kf.
8. (hip adj2 resurfacing*).tw,kf.
9. ((resurfacing adj2 arthroplasty) and hip).tw,kf.
10. (hip adj2 prosthesis*).tw,kf.
11. (stem* or cup or cups or acetabul*).m_titl.
12. randomized controlled trial.pt.
13. (random* or blind*).tw,kf.
14. (systematic review or meta-analysis).pt.
15. (systematic review or meta-analysis).tw,kf.
16. 12 or 13 or 14 or 15
17. 1 or 2 or 3 or 4
18. 5 or 6 or 7 or 8 or 9 or 10 or 11
19. 16 and 17 and 18
20. limit 19 to english language
21. limit 20 to yr="2000 -Current"
22. (comment or letter or editorial).pt.
23. 21 not 22

List of sources to be searched:

- Full-text publications in MEDLINE, EMBASE, Cochrane Library back to inception
- Meeting abstracts back to 2017 in EMBASE
- HTA database www.crd.york.ac.uk/PanHTA/
- NICE www.nice.org.uk/
- INAHTA www.inahta.org/Publications/
- SIGN www.sign.ac.uk/
- Canadian Agency for Drugs and Technologies in Health (CADTH) www.cadth.ca/
- IQWiG www.iqwig.de/en/methods/results/hta-reports.7459.html
- Research in progress: ClinicalTrials.gov <https://clinicaltrials.gov/> and WHO Clinical Trials Search Portal <http://apps.who.int/trialsearch/> and NIHR Awards <https://fundingawards.nihr.ac.uk/search>

Appendix 2

Risk-of-bias analysis

Risk of bias was assessed in the RCTs that met the criteria for inclusion in the study.

Chapter 3: non-regenerative procedures used to treat early osteoarthritis in the knee

1. Risk of bias in RCTs (Figure 23)

Randomisation bias

We assessed two RCTs to be at high risk of randomisation bias, one because of probable inadequate randomisation processes and one because of a large difference in outcome at baseline that was larger than the effect size.^{81,97}

	Randomisation bias	Deviations from interventions	Missing data	Outcome measurement	Selection of reported result	Overall risk of bias
Bisson <i>et al.</i> ⁸⁰	+	+	?	+	+	?
Chang <i>et al.</i> ⁸¹	–	+	+	–	?	–
Forster <i>et al.</i> ⁸²	?	–	+	–	?	–
Gauffin <i>et al.</i> ⁸³	+	–	?	–	+	–
Herrlin <i>et al.</i> ⁸⁴	+	+	+	–	?	–
Hubbard <i>et al.</i> ⁸⁵	+	+	+	?	?	?
Kalunian <i>et al.</i> ⁸⁶	?	+	+	+	?	?
Katz <i>et al.</i> ⁸⁷	+	–	?	–	+	–
Kirkley <i>et al.</i> ⁸⁸	+	+	?	–	?	–
Kise <i>et al.</i> ⁸⁹	?	–	?	–	+	–
Merchan <i>et al.</i> ⁹⁰	+	+	+	–	?	–
Moseley <i>et al.</i> ²¹	+	+	?	?	?	–
Moseley <i>et al.</i> ²¹	+	+	?	?	?	–
Roos <i>et al.</i> ⁹¹	+	–	?	+	+	–
Saeed <i>et al.</i> ⁹²	?	?	+	–	?	–
Sihvonen <i>et al.</i> ²²	?	+	+	+	+	?
Spahn <i>et al.</i> ⁹³	+	+	?	?	+	?
van der Graaf <i>et al.</i> ⁹⁵	+	–	?	–	+	–
Vermesan <i>et al.</i> ⁹⁶	?	–	+	–	?	–
Yin <i>et al.</i> ⁹⁷	–	+	+	–	?	–

FIGURE 23 Risk of bias in RCTs for non-regenerative procedures used to treat early knee OA.

Deviation from intended interventions bias

We assessed seven RCTs to be at high risk of deviations from intended interventions bias because differential numbers of participants in each group switched treatment arms (mainly switching from the non-operative arm to the surgical arm); this switching was known to the trial personnel and participants and was likely to affect the outcome.^{82,83,87,89,91,95,96}

Outcome measurement (reporting bias)

We assessed 12 RCTs to be at high risk of outcome measurement bias for the PROMS outcomes because the outcome assessors (participants) were not blinded to the treatment and were likely to have been influenced by that knowledge.^{81–84,87–90,92,95–97}

Missing data (attrition bias) and selective reporting bias

We considered no RCTs to have a high risk of missing data bias or selective reporting bias for all time points. One RCT rated as 'some concerns' had no loss to follow-up at short-term (1 year), but by medium-term (4 years), 33% of patients (13% of the intervention group and 47 % control) were not included in PROMS analyses as they underwent further surgical procedures. By 10-year follow-up, only 37% of the initial cohort was analysed for PROMS (missing data: 57% intervention and 70% control), thus at short-term, there was low risk of attrition bias, but at long term, high risk of bias.

Overall risk of bias

The overall risk of bias for five RCTs was 'some concerns'.^{22,80,85,86,93} All other RCTs were assessed as high risk of bias (see [Figure 23](#)). We considered no RCTs to have low risk of bias for all domains.

2. Risk of bias in non-randomised comparative studies for non-regenerative procedures used to treat early knee OA ([Table 37](#))

We assessed study quality using the NIH tool for cohort studies. We considered one study NRS included to be at overall high risk of bias, determined particularly by the domain of selection bias (more than 50% refused to take part in the retrospective analysis, participants who refused surgery were assigned to the non-operative management group), and the lack of adjustment for confounders for the TKA outcome.⁹⁴

Chapter 4: non-regenerative procedures used to treat early osteoarthritis in the hip

1. Risk of bias in RCTs ([Figure 24](#))

Randomisation bias

All studies were rated as low for randomisation bias as they reported robust methods for allocation concealment with either adequate or no information on sequence generation and baseline imbalance.

Deviation from intended interventions bias

We assessed one study as high risk of deviations from intended interventions bias due the substantial proportion (70%) of patients who switched from non-operative treatment to the surgical intervention arm.¹⁰⁴

Missing data (attrition bias) and Selective reporting bias

Three studies were assessed as having some concerns in regard to missing data.^{104–106} In one study the risk of attrition bias was rated 'some concerns' for the PROMs outcomes and low for reoperations.¹⁰⁶ In two of these studies, the proportion of missing data were unbalanced, with a higher proportion in the non-operative treatment group.^{104,105}

Outcome measurement (reporting bias)

Three RCTs comparing non-operative to operative treatments were rated to be at high risk of bias for the PROMS outcomes because the outcome assessors (participants) were not blinded to the treatment and were likely to have been influenced by that knowledge.^{103–105}

TABLE 37 Risk of bias in non-randomised comparative studies for non-regenerative procedures used to treat early knee OA

Criteria	Y/N/CD/NR or N/A
Study	Su et al.⁹⁴
1. Research question/objective clearly stated?	Y
2. Selection bias 1: study population clearly specified and defined?	Y
3. Selection bias 2: participation rate of eligible persons at least 50%?	N
4. Selection bias 3: all participants selected/recruited from same/similar populations?	N
5. Sample size justification, power description, or variance and effect estimates provided?	NR
6. Intervention measured prior to the outcome(s) being measured?	Y
7. Time frame sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed?	Y
8. Study examined different levels of the exposure as related to outcome?	N/A
9. Interventions clearly defined, valid, reliable and implemented consistently across all study participants?	Y
10. Exposure(s) assessed more than once over time?	N/A
11. Outcome measurement bias 1: outcome measures clearly defined, valid, reliable and implemented consistently across all study participants?	Y
12. Outcome measurement bias 2: outcome assessors blinded to the exposure status (intervention) of participants?	No for WOMAC; CD for TKA
13. Attrition bias: was loss to follow-up after baseline 20% or less?	Y
14. Key confounders measured and adjusted statistically for their impact on the relationship between intervention and outcome?	Y (partly for WOMAC); N for TKA (despite stratification by KL grade)
Overall risk of bias (low, moderate, high).	High
CD, cannot determine; N, no; N/A not applicable; NR, not reported; Y, yes.	

	Randomisation bias	Deviations from interventions	Missing data	Outcome measurement	Selection of reported result	Overall risk of bias
Ayeni et al. ¹⁰⁶	+	+	?	+	+	?
Griffin et al. ¹⁰³	+	?	+	-	+	-
Krych et al. ¹⁰²	+	+	+	?	?	?
Mansell et al. ¹⁰⁴	+	-	?	-	+	-
Palmer et al. ¹⁰⁵	+	+	?	-	+	-

FIGURE 24 Risk of bias in RCTs of non-regenerative procedures used to treat early hip OA.

Overall risk of bias

Overall risk of bias was rated as high in the three studies comparing a surgical procedure with conservative treatment.¹⁰³⁻¹⁰⁵ The two studies comparing two different surgical procedures had overall risk of bias of some concerns; in one of these the overall risk of bias was low for the outcome of reoperations.^{102,106}

2. Risk of bias in NRS ([Table 38](#))

We assessed the one NRS selected using the NIH tool for cohort studies.¹⁰⁷ The study was rated as low/moderate risk of bias of overall bias. This study had a clearly defined patient population identified with CTP codes (labral tears) and used propensity score matching for age, sex, comorbidities, ethnicity and smoking status.

Chapter 5: load-modifying procedures used to treat early osteoarthritis in the knee.

1. Chapter 5 – Risk of bias in RCTs ([Figure 25](#))

Randomisation bias

We assessed the van der Woude study to be high risk of randomisation bias because the difference in baseline PROMs was larger than the effect estimate.¹¹³

Deviation from intended interventions bias

There was no evidence of deviation from the intended intervention in any of the studies.

TABLE 38 Risk of bias in non-randomised comparative studies of non-regenerative procedures used to treat early hip OA

Criteria: Y/N/CD/NR or N/A	Scott <i>et al.</i> ¹⁰⁷
1. Research question/objective clearly stated?	Y
2. Selection bias 1: study population clearly specified and defined?	Y
3. Selection bias 2: participation rate of eligible persons at least 50%?	CD
4. Selection bias 3: all participants selected/recruited from same/similar populations?	Y
5. Sample size justification, power description or variance and effect estimates provided?	Y
6. Exposure(s) of interest measured prior to the outcome(s) being measured?	Y
7. Time frame sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed?	Y
8. Did study examine different levels of the exposure as related to the outcome?	N/A
9. Classification of interventions bias: interventions clearly defined, valid, reliable and implemented consistently across all study participants?	Y
10. Exposure(s) assessed more than once over time?	N/A
11. Outcome measurement bias 1: outcome measures clearly defined, valid, reliable and implemented consistently across all study participants?	Y
12. Outcome measurement bias 2: outcome assessors blinded to the exposure status (intervention) of participants?	N/A
13. Attrition bias: was loss to follow-up after baseline 20% or less?	NR
14. Confounding bias: were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?	Y
Overall risk of bias (Low = L; Moderate = M; High = H).	L/M

CD, cannot determine; N, no; N/A not applicable; NR, not reported; Y, yes.

	Randomisation bias	Deviations from interventions	Missing data	Outcome measurement	Selection of reported result	Overall risk of bias
Shaofei <i>et al.</i> ¹¹²	?	+	+	?	?	?
van de Woude <i>et al.</i> ¹¹³	-	+	-	?	+	-
Brouer <i>et al.</i> ¹¹⁶	?	+	-	?	-	-

FIGURE 25 Risk of bias in RCTs involving load-modifying procedures used to treat early knee OA.

Missing data (attrition bias) and selective reporting bias

In the van der Woude study there was a high risk of missing data bias at 2 years because 9% of the participants receiving KJD dropped out because of ‘unsatisfactory treatment benefit’ (and 9% of HTO participants were lost to follow-up because of comorbidities unrelated to the treatment).¹¹³ The Brouwer study had high risk for missing data as they omitted preoperative PROMS scores and did not report on the completeness of postoperative collection.¹¹¹

Outcome measurement (reporting bias)

As noted above in the Brouwer study there was no reporting of preoperative reporting PROM scores.¹¹¹

Overall risk of bias

The overall risk of bias was assessed to be high for both the study comparing HTO and KJD and the study comparing two different types HTO.^{111,113} Risk of bias was assessed as ‘some concerns’ for the RCT comparing HTO and UKA.¹¹²

2. Chapter 5 – Risk of bias in NRS (Table 39)

We assessed study quality using the NIH tool for one NRS. We assessed overall risk of bias as moderate for the van Outeren 2017 study for the comparison of HTO versus NICE core + brace, and moderate/high for the HTO versus NICE core comparison (UC).¹¹⁴ For this study, we had concerns about selection into the study for the comparator group: propensity score matching resulted in only 30 brace participants (from a pool of 60, 12 of whom were excluded because of valgus malalignment – because the HTO participants only had varus malalignment) versus 83 HTO (from 92) patients; and in the other comparison, 28 UC participants (from a pool of 57; 10 of whom had valgus malalignment) versus 71 HTO (from 92) participants. It is unclear whether the 18/48 (38%) non-matched brace patients and the 19/47 (40%) non-matched UC patients corresponded to the 25 and 14 participants described in the original brace versus UC trial, who discontinued the original interventions, mostly because of ineffectiveness of the intervention. If so, this might underestimate the effect of the surgery versus non-operative management. Missing data in the van Outeren 2017 NRS were generally low for the HTO versus brace comparison: 1/83 (1.4%) HTO and 3/10 (10%) brace – this may imply those not matched were patients who did not adhere to treatments. In the comparison of HTO versus UC, there were 6/28 (21%) participants with missing data in the UC group and only 1/71 (1%) missing for HTO, which is an additional risk of bias.

TABLE 39 Risk of bias in non-randomised comparative studies involving load-modifying procedures used to treat early knee OA

Criteria	Y/N/CD/NR or N/A
	Van Outeren <i>et al.</i> ¹¹⁴
1. Research question/objective clearly stated?	Y
2. Selection bias 1: study population clearly specified and defined?	Y
3. Selection bias 2: participation rate of eligible persons at least 50%?	NR
4. Selection bias 3: all participants selected or recruited from the same/similar populations?	CD
5. Sample size justification, power description or variance and effect estimates provided?	N
6. Exposure(s) of interest measured prior to the outcome(s) being measured?	Y
7. Time frame sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed?	Y
8. Did study examine different levels of the exposure as related to the outcome?	N/A
9. Classification of interventions bias: interventions clearly defined, valid, reliable and implemented consistently across all study participants?	Y
10. Exposure(s) assessed more than once over time?	N/A
11. Outcome measurement bias 1: outcome measures clearly defined, valid, reliable and implemented consistently across all study participants?	Y
12. Outcome measurement bias 2: outcome assessors blinded to the exposure status (intervention) of participants?	CD
13. Attrition bias: was lost to follow-up after baseline 20% or less?	Y
14. Confounding bias: were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?	Y
Overall risk of bias	Moderate for HTO vs. brace and moderate/high for HTO vs. UC
CD, cannot determine; N, no; N/A not applicable; NR, not reported; Y, yes.	

EME
HSDR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

*This report presents independent research funded by the National Institute for Health and Care Research (NIHR).
The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the
Department of Health and Social Care*

Published by the NIHR Journals Library